Dynamic causal model diagnostic reasoning for online technical process supervision

Jacky Montmain, Sylviane Gentil

Dynamic causal model diagnostic reasoning for online technical process supervision☆

Jacky Montmain[a], Sylviane Gentil[b,*]

[a]LGI2P, URC EMA-CEA, Site EERIE, 30035 Nîmes Cedex 1, France
[b]CNRS-INPG-UJF, Laboratoire d'Automatique de Grenoble, BP 46, 38402 Saint Martin d'Hères Cedex, France

Abstract

Model-based diagnosis is founded on the construction of fault indicators. The methods proposed for this purpose generally represent the process by means of an extremely inflexible formalism that limits the scope of applications. Moreover, it is usually difficult and costly to develop precise mathematical models of complex plants. New and more flexible techniques intended notably to explain the observed behavior open new perspectives for fault detection and diagnosis. The diagnostic procedures for such plants are generally integrated into a supervisory system, and must therefore be provided with explanatory features that are essential interpretation and decision-making supports. Techniques based on causal graphs constitute a promising approach for this purpose. A causal graph represents the process at a high level of abstraction, and may be adapted to a variety of modeling knowledge corresponding to different degrees of precision in the underlying mathematical models. When the process is dynamic the causal structure must allow temporal reasoning. Lastly, because reasoning on real numbers is often used by human beings, fuzzy logic is introduced as a numeric–symbolic interface between the quantitative fault indicators and the symbolic diagnostic reasoning on them; it also provides an effective decision-making tool in imprecise or uncertain environments. An industrial application in the nuclear fuel reprocessing industry is presented.

Keywords: Supervision; Fault isolation; Fault filtering; Causal reasoning; Diagnostic inference; Fuzzy inference

1. Introduction

An operator support system that makes it possible to assess the process state at all times, to diagnose failures and to supply validated data to the process control system would be an attractive way to improve large-scale plant availability and maintainability and more generally plant dependability. These considerations explain the rapid development of plant supervision and the success encountered by process fault detection and isolation (FDI) techniques in recent years (Gertler, 1997, 1998; Chen, Patton & Zhang, 1996). However, in the light of the recent analysis proposed by Isermann and Ballé (1997), much work remains to be done before the modeling formalisms and methods proposed in the literature meet the general objective described above and can actually be applied to industrial processes.

There are a number of reasons for this situation, including the inadequacy of available models for the proposed FDI algorithms, the difficulty of compiling the diverse knowledge required for plant diagnosis, and the lack of instrumentation designed specifically for diagnosis.

Two types of models are generally available for industrial plants: material or energy balances established from process block diagrams and flowsheets that integrate operator knowledge of production rules, and complex, partial derivative nonlinear analytical equations that are written by physicists or chemists when the focus is the process and the physical phenomena involved. In large plants, the former are written from a production management standpoint and thus implement shop-scale balances, while the latter are developed to obtain load diagrams or to build training simulators. In any case,

☆This paper was not presented at any IFAC meeting. This paper was recommended for publication in revised form by Associate Editor T.A. Johansen under the direction of Editor S. Skogestad.

* Corresponding author. Tel.: (+ 33)-4-76-82-62-39; fax: (+ 33)-4-76-82-63-88.

E-mail addresses: jacky.montmain@site-eerie.ema.fr (J. Montmain), sylviane.gentil@lag.ensieg.inpg.fr (S. Gentil).

they were conceived for purposes other than supervision. Classic FDI techniques used in automatic control — generalized parity space, dedicated observers scheme or parameter estimation (Frank, 1990, 1991; Patton & Chen, 1991; Isermann, 1993) — are poorly suited to this type of representation. Classic process diagnosis techniques which are generally based on state variable representation, are not adapted to the supervision of a complete facility because of their constraining formalism and global analytical processing.

In an attempt to surmount these shortcomings, new techniques (notably artificial intelligence and qualitative modeling) have been adopted in recent years (Frank, 1994, 1996). But propagating qualitative values generally leads to multiple solutions, which is not compatible with diagnosis (MQ&D coordinated by P. Dague, 1995). Some attempts to overcome this issue have been proposed in Montmain, Leyval and Gentil (1994a), and Mosterman, Biswas and Manders (1998).

The *hybrid method* presented here relies on both a qualitative causal representation of the process function and quantitative local behavioral models. It allows the construction of a complete FDI system for less constraining representations than those of process control theory. Moreover, any existing models of the industrial facility that were initially designed for integration in training simulators can be reused without further refining. It generalizes previous results to a large class of numerical models (Montmain & Gentil, 1993; Penalva, Coudouneau, Leyval & Montmain, 1993).

Another reason for the choice of a hybrid approach is not a technical one and is much less often discussed in the literature; it concerns the very purpose of the supervision system in the control room. Until now, measurement consistency analysis and fault detection functions were implicitly performed by human operators; thus they were generally overlooked or only partially incorporated in the automation scheme.

One point of view is to consider that the objective of supervision is to automate these decision-making tasks. An *ideal* supervisory control and data acquisition system (SCADA) should be able to reconfigure and tune new control laws when a degraded operating mode is detected. The operator is removed from the control loop in this *active supervision* perspective. From this standpoint, the objective is not to help the operator in a critical situation to analyze a process that computerization has made increasingly complex, but rather to relieve him of decision-making tasks in a faulty situation.

Another point of view is to consider that the exhaustiveness of the mathematical model in such an approach is illusive and consequently that FDI techniques are not ready to be extended to the supervision of plant facilities. As a result, the recovery action in the event of a malfunction in a complex large-scale process will remain subject to the decision of the operating staff for many years to come. The objective of supervision is thus revised: it is more reasonable to envisage systems that no longer eliminate the operators from process control, but instead support them in this function by assisting their reasoning on line (Rasmussen, 1993; Montmain, 1997).

A supervision system that targets the operating team rather than on the SCADA leads to major revisions in model-based diagnosis strategies (Montmain, 1998). The intelligibility and pertinence to the operator of the results provided by a diagnostic system become legitimate issues. The results of the diagnostic system become part of the reasoning of the operator analyzing the situation and as such must be substantiated and explained.

The approach discussed here is based on this second *passive supervision* option and relies on causal and approximate reasoning on quantitative data, issued from on line sensors and numerical models. This is a more natural form of reasoning for human beings. As a result, it is easily interpreted by an operator.

Additionally, it is not always sufficient to isolate the primary fault in operating situations. In many large industrial plants, the operators benefit from an appreciable control margin with respect to the nominal operating parameters. Isolation is, of course, vital for a maintenance policy or when the fault requires an immediate reaction. However, if the operators consider that the available control margin makes an immediate reaction unnecessary on detection of the fault, they will wait for further and more significant events before undertaking a counter-action. A dynamic monitoring of the primary fault effects is necessary to ensure a continuous assessment of the disturbed functions and to revise the initial decision if necessary (Mosterman, Biswas & Narasimham, 1997). This concept of fault filtering is generally disregarded in classic diagnostic systems where the objectives are early detection and isolation on the basis of a theoretical fault signature. From this standpoint, the diagnosis of an industrial process differs considerably from theoretical diagnosis: supervision requires filtering new faults and relating them if appropriate to the problem identified earlier rather than minimal response time FDI.

Section 2 describes the diagnostic enhancements of causal reasoning, relates the proposed approach to other causal model-based approaches and shows how complex plants can be represented by a causal structure. The numerical model implemented with this causal structure can take practically any form. The first attempt at dynamic causal reasoning was initiated by the particular model proposed by Leyval, Gentil and Feray-Beaumont (1994), and its simulation, explanatory and advisory capabilities have been enhanced. This article extends the notion of causality to other types of models and demonstrates its utility for diagnosis through detection, isolation and alarm filtering procedures. Section 3 is dedicated to residual generation. The models used to

generate residuals are numerical ones. The techniques of generating and processing residuals locally for a variable are discussed, together with overall control of diagnostic reasoning at the level of the causal structure. Section 4 reviews the advantages of using fuzzy reasoning to model FDI and fault filtering as decision making processes in an imprecise environment due to modeling imprecision and measurement uncertainty. Section 5 examines the industrial application of these techniques to nuclear fuel reprocessing using a causal graph modeling the normal dynamic behavior of the facility. A glossary and an appendix complement the paper.

2. Causal graph-based diagnosis

Diagnosis is typically a causal process for it consists in pointing out the faulty components that can explain the observed malfunction. Davis (1983) wrote that a significant aspect of the knowledge required to analyze disturbed regimes is an understanding of the mechanisms in causality terms. A causal structure is a description of the effects that variables may have on one another, and it may be represented by a directed graph (digraph). This structure then provides a conceptual tool for reasoning about the way in which normal or abnormal changes propagate within a plant. The nodes are the variables and the arcs symbolize the relations among them.

All causal graph-based diagnostic methods implement the same basic principle. The objective is to account for deviations detected in the evolution of the variables of a plant using a minimum of malfunctions at the source. If significant deviations are detected, primary faults are hypothesized and the propagation paths in the directed graph are analyzed to determine whether this failure hypothesis is sufficient to account for the remaining faults. A primary fault is a change in the evolution of a variable that is directly attributable to a failure or to an unmeasured disturbance; secondary faults result from the propagation of this change in the process over time, causing new deviations. Causal graph-based diagnosis consists in finding the source variable whose variation accounts for all the deviations detected on the other variables (the detection variables) (Montmain & Leyval, 1994b). The algorithm for locating the primary deviation is a backward/forward procedure starting from a detected variable: the backward search formulates hypotheses, and the forward search evaluates them. The backward search bounds the fault space by eliminating the normal measurements causally upstream. Then each possible primary deviation generates a hypothesis which is forward tested by using the states of the variables and the functions of the arcs.

As an example, the simplest causal graph structure is the signed digraph (SDG). A diagnostic method using an SDG as the basic data structure was initially presented by Iri, Aoki, O'Shima and Matsuyama (1980). The nodes of the SDG correspond to variables, and the directed branches are labeled by signs: the sign is defined as " + " when the variables of the arc evolve in the same direction, and the sign is " − " when they evolve in the opposite direction. The state of a variable is expressed in the quantity-space $\{+, 0, -\}$, according to whether the value is normal (0), higher than normal (+), or lower than normal (−). The graph exclusively composed of signed nodes (+ or −) and consistent branches — branches for which the product of the signs of the initial and the final nodes is the same as the sign of the branch — is a representation of the propagation of the fault in the system. This initial and final nodes *consistency test* is recursively carried out and constitutes the basic isolation procedure. The roots of such a sub-graph are candidates for the origins of the failure.

Shiozaki, Matsuyama, Tano and O'Shima (1985) conclude that the purely qualitative description of variables is too rough and explain that the diagnosis results cannot be accurate if there is a poor balance between the thresholds of the variables which are connected by an arc. As a solution to this issue, an algorithm based on a five-range pattern of the variable states $\{-, -?, 0, +?, +\}$ is proposed to avoid the pitfalls of a wrong diagnosis. Ambiguity symbols (− ? or + ?) associated with the nodes improve the robustness of even bad thresholds choices. The algorithm proposed by Palowitch and Kramer (1986) also uses the simplest causal graph; the arcs are labeled by signs, but numerical information is stored in nodes: their *deviation indexes* DI are a normalized measure of the deviation from the steady state.

Yu and Lee (1991) apply fuzzy sets to manage progressive quantification of the qualitative digraph structure. The membership function of the fuzzy set theory provides a simple way to integrate quantitative knowledge in qualitative representations.

Finally, in the method presented here, the behavior of variables is purely numerically described (measurements, estimations and associated residuals) and the functions symbolized by the arcs are numerical differential equations. Fuzzy set theory is introduced at a higher level of abstraction to symbolically interpret and manage the numerical residuals in a diagnosis reasoning as explained in Section 4.

A second point is that the process is dynamic; therefore the signatures of the observed faults change over time which implies that temporal fault filtering is a required functionality. Montmain and Gentil (1993) introduce temporal information within the arcs. In Mosterman et al. (1997) the idea is to predict future behavior of the system for each abnormal deviation in terms of their qualitative time-derivative changes. After initial component parameter implication and prediction (backward and forward chaining), the progressive monitoring module compares reported signatures and actual observations as

they change dynamically after faults have occurred. Progressive monitoring is activated when there is a discrepancy between a predicted value and a monitored value that deviates. At every time point, it is determined whether the next higher derivative could make the prediction consistent with the observation.

The fault filtering method presented here could be assimilated with an extension of the progressive monitoring: a measured variable is no longer described by its qualitative value and other higher-order derivatives but through the dynamic equation that manages the residual behavior associated with it. Numerical residuals make it possible to distinguish normal dynamic effects from fault propagation.

In conclusion, the causal structure represents the process at a high level of abstraction. It can support a great variety of information. The digraph is above all a reasoning structure that can be enriched as knowledge becomes available. This paper deals with *the quantitative dynamic case in causal graph-based diagnosis*.

An example of a dynamic causal graph managing quantitative and temporal information in an event-driven causal simulator is described in Leyval et al. (1994). This graph was obtained by a careful physical analysis and a functional top-down breakdown of the process. Nodes were selected as variables meaningful for the supervision operator. It particularly focused on hydraulic phenomena (balances, transfers, storage, etc.). Temporal parameters in the dynamic relations supported by the arcs were obtained with classical identification procedures (Leyval & Ledoux, 1991).

In the case of a process for which a model described by a system of differential equations is available, the causal relations among the variables are implicit: the effects on the output cannot precede the input variations. Defined in this way, causality becomes equivalent to the notion of calculability, and is related to the implicit sampling of differential equations by the simulation algorithms. This means that the differential equation system can be represented by a causal graph that may include loops.

$$\begin{aligned}
\dot{x}_1 &= g_1(u_1, u_n, x_1),\\
\dot{x}_2 &= g_2(u_i, x_2),\\
\dot{x}_3 &= g_3(u_1, x_2, x_n, x_3)\\
&\vdots\\
\dot{x}_n &= g_n(x_3, x_4, \ldots, x_n).
\end{aligned} \tag{1}$$

The output of one equation becomes the input of the following equation, corresponding to the choice of a partial strict order relation as shown in Fig. 1.

If it is initially assumed that the g_i relations are linear, and that X_i is a process variable, then based on the notation in Fig. 1, the simulator computes the behavior

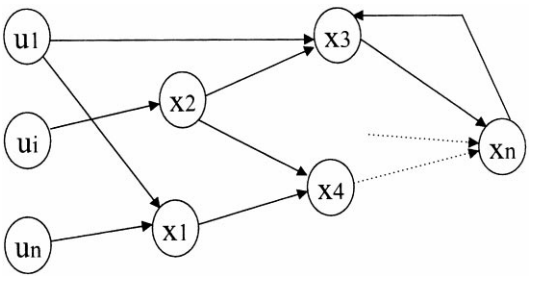


Fig. 1. Causality: A partial strict-order relation.

of X_i as follows:

$$x_i(k) = (1 - A_i(q^{-1}))x_i(k) + \sum_{j\in P_i} q^{-d_{ji}} A_{ji}(q^{-1})x_j(k) + \sum_{j\in U_i} q^{-d'_{ji}} B_{ji}(q^{-1})u_j(k), \tag{2}$$

where A_i, A_{ji}, B_{ji} are polynomials in the shift operator q and d_{ji}, d'_{ji} represent the delays with respect to other variables or inputs; x_j are the computed behavior of the predecessors X_j, P_i designates the set of all subscripts j of the predecessors X_j of X_i, and U_i refers to the set of subscripts j of the process inputs directly affecting X_i.

3. Generation of residuals with the causal structure

This section proposes a method for managing residuals based on the causal graph; thus it provides the isolation power of numerical methods without the need for algebraic operations to obtain a structured residual set. It allows for a wider range of process representations than those proposed by classic diagnostic methods; notably it is flexible enough for possible application to industrial simulators (Section 1) with no additional modeling. In view of this orientation, the following discussion is illustrated by a discrete representation for the differential equations.

3.1. Residual generation

We assume initially that the model is accurate enough to give the actual behavior $\bar{x}_i$ of X_i by the difference equation:

$$\bar{x}_i(k) = (1 - A_i(q^{-1}))\bar{x}_i(k) + \sum_{j\in P_i} q^{-d_{ji}} A_{ji}(q^{-1})\bar{x}_j(k) + \sum_{j\in U_i} q^{-d'_{ji}} B_{ji}(q^{-1})u_j(k) + E_i(q^{-1})f_i(k), \tag{3}$$

where f_i is the variable modeling the effect of failures liable to affect X_i directly.

The simulator outputs are compared with the process sensor outputs. The error between the measured and simulated values constitutes the open-loop residual ε_i:

$$\varepsilon_i(k) = \bar{x}_i(k) - x_i(k). \tag{4}$$

The dynamics of ε_i are given by the following relation:

$$\varepsilon_i(k) = (1 - A_i(q^{-1}))\varepsilon_i(k) + \sum_{j \in P_i} q^{-d_{ji}} A_{ji}(q^{-1})\varepsilon_j(k) + E_i(q^{-1})f_i(k). \tag{5}$$

These residuals alone permit only detection: it is clear that ε_i is excited either by a local fault f_i or by a fault affecting an upstream variable X_j. Since the idea is that the use of a model other than the simulator is to be avoided for isolation purposes, the diagnostic system must be provided with another set of fault indicators using the potential inputs to the simulator.

At this stage in the modeling procedure, the dynamics of the simulation residuals are expressed locally while taking advantage of the causal structure of the system. Another type of residual — called prediction residual — is based on the preceding causal decomposition and uses no parameters other than those of the simulator equations. Consider the following equation:

$$x_i^p(k) = (1 - A_i(q^{-1}))x_i^p(k) + \sum_{\substack{j \in P_i \\ j \neq p}} q^{-d_{ji}} A_{ji}(q^{-1})x_j(k) + q^{-d_{jp}} A_{jp}(q^{-1})\bar{x}_p(k) + \sum_{j \in U_i} q^{-d'_{ji}} B_{ji}(q^{-1})u_j(k). \tag{6}$$

In this equation, the simulated evolution x_p of X_p has been replaced by its measured evolution $\bar{x}_p$ to obtain the predicted evolution x_i^p of X_i. This will be referred to as predicting X_i by reconfiguring the predecessor X_p.

It is now possible to define the prediction residual after reconfiguring X_p by subtracting Eq. (6) from Eq. (3):

$$\varepsilon_i^p(k) = \bar{x}_i(k) - x_i^p(k), \tag{7}$$

$$\varepsilon_i^p(k) = (1 - A_i(q^{-1}))\varepsilon_i^p(k) + \sum_{\substack{j \in P_i \\ j \neq p}} q^{-d_{ji}} A_{ji}(q^{-1})\varepsilon_j(k) + E_i(q^{-1})f_i(k). \tag{8}$$

From a more general standpoint, several predecessors may be reconfigured at the same time. If P_i^r is the set of subscripts of the reconfigured predecessors, then $\varepsilon_i^{P_i^r}$ designates the prediction error on reconfiguration of the predecessors associated with P_i^r:

$$\varepsilon_i^{P_i^r}(k) = \bar{x}_i(k) - \bar{x}_i^{P_i^r}(k), \tag{9}$$

$$\varepsilon_i^{P_i^r}(k) = (1 - A_i(q^{-1}))\varepsilon_i^{P_i^r}(k) + \sum_{\substack{j \in P_i \\ j \notin P_i^r}} q^{-d_{ji}} A_{ji}(q^{-1})\varepsilon_j(k) + E_i(q^{-1})f_i(k). \tag{10}$$

In a linear system, it can easily be shown that $\varepsilon_i^{P_i^r}$ may be calculated from ε_i^p, $p \in P_i^r$: individual reconfigurations provide access to all the prediction errors obtained on combinations of reconfigurations. The limit case is the prediction error, $\varepsilon_i^{P_i}$: the calculated evolution of X_i is obtained only from measurements of its predecessors.

Finally, subtracting Eq. (8) from Eq. (5) yields

$$\varepsilon_i(k) - \varepsilon_i^p(k) = (1 - A_i(q^{-1}))(\varepsilon_i(k) - \varepsilon_i^p(k)) + q^{-d_{pi}} A_{pi}(q^{-1})\varepsilon_p(k), \tag{11}$$

which may be written as

$$\varepsilon_i(k) - \varepsilon_i^p(k) = \frac{q^{-d_{pi}} A_{pi}(q^{-1})}{A_i(q^{-1})}\varepsilon_p(k). \tag{12}$$

In the case of a multiple reconfiguration, Eq. (12) becomes

$$\varepsilon_i(k) - \varepsilon_i^{P_i^r}(k) = \sum_{p \in P_i^r} \frac{q^{-d_{pi}} A_{pi}(q^{-1})}{A_i(q^{-1})}\varepsilon_p(k). \tag{13}$$

Eq. (12) implies that the contribution of the simulation errors of the selected predecessors can be determined simply by comparing the suitable prediction error ε_i^p with the simulation error ε_i. The difference between the simulation error and the prediction error on reconfiguring X_p can be used to assess the contribution of the simulation error ε_p propagated from X_p to the observed error on X_i.

Fig. 2 reviews the excitation sources for Eqs. (12), (5) and (8).

The open loop residual Eq. (4) is used for detection.

The next section develops the isolation procedure using the prediction errors Eqs. (12) and (13).

3.2. Fault isolation and filtering

For a variable X_i, it is possible to use the prediction residuals with different reconfigurations to determine whether the observed error on X_i has a local cause — primary fault — or is simply the consequence of an upstream fault — secondary fault. This test is the basic isolation and fault filtering procedure. The proposed method estimates and quantitatively analyzes the simulation and prediction errors to determine whether the errors are related. Establishing a link between two deviations is not based only on the sign of the arcs as in a simple SDG test, but also involves a dynamic quantitative analysis of the local model involved in the arc.

Consider the simple example in Fig. 3. The simulated evolution of X_3 is calculated with the simulated evolutions of X_1 and X_2; X_3 is the detection variable. A significant error ε_3 is detected on X_3, and the problem is to

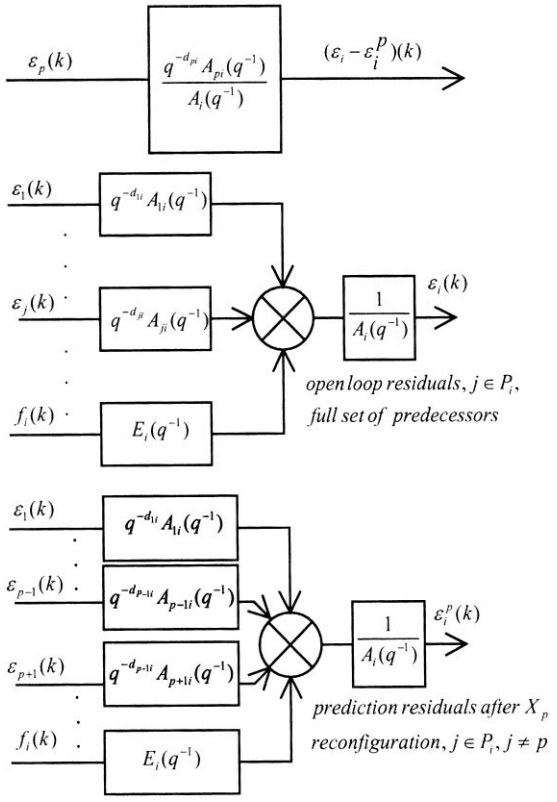


Fig. 2. Open-loop and prediction residuals.

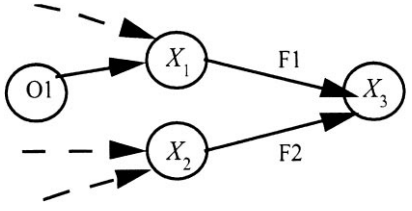


Fig. 3. A causal graph.

identify its origin: it may be due to a failure f_3 directly affecting X_3, or may simply be a consequence of a fault f_1 affecting X_1 or a fault f_2 affecting X_2. Local reconfigurations are used to test these conditions. A local prediction in which the actual evolution of X_1 is substituted for its simulated evolution is performed as if X_1 were a simulator input. The same is done with X_2.

Reconfiguring X_1 and/or X_2 yields three prediction errors: ε_3^1 and ε_3^2 (Eq. (12)) and $\varepsilon_3^{\{1,2\}}$ (Eq. (13)). According to Fig. 2, Eqs. (12) and (13), the following cases may be noted:

- If ε_3^j and ε_3 are nearly identical, hereafter denoted by $\varepsilon_3^j \cong \varepsilon_3$, then $\varepsilon_j \cong 0$ and X_j cannot be considered as responsible for the fault detected on X_3;
- If ε_3^j is negligible compared with ε_3, hereafter denoted by $\varepsilon_3^j \ll \varepsilon_3$, then ε_3 is essentially excited by ε_j according to Eq. (12), and X_j is incriminated as responsible for ε_3.
- If ε_3^j and ε_3 are of the same order of magnitude, hereafter denoted by $\varepsilon_3^j \approx \varepsilon_3$, it may be concluded merely that ε_3 is only partially explained by ε_j. In this case, Eq. (13) can account for ε_3 by multiple faults. In addition, if $\varepsilon_3^{\{1,2\}} \ll \varepsilon_3$, both X_1 and X_2 are incriminated.

This can be summed up as follows:

$$\begin{aligned}
&\varepsilon_3^1 \ll \varepsilon_3 \text{ and } \varepsilon_3^2 \cong \varepsilon_3\text{: secondary fault incriminating } X_1,\\
&\varepsilon_3^1 \cong \varepsilon_3 \text{ and } \varepsilon_3^2 \ll \varepsilon_3\text{: secondary fault incriminating } X_2,\\
&\varepsilon_3^1 \cong \varepsilon_3 \text{ and } \varepsilon_3^2 \cong \varepsilon_3\text{: primary fault incriminating } X_3,\\
&\varepsilon_3^1 \approx \varepsilon_3 \text{ and } \varepsilon_3^2 \approx \varepsilon_3 \text{ and } \varepsilon_3^{\{1,2\}} \ll \varepsilon_3\text{: secondary fault}\\
&\quad\text{incriminating } X_1 \text{ and } X_2.
\end{aligned} \tag{14}$$

Fig. 2 illustrates that structured residuals for fault isolation are obtained without any algebraic manipulation thanks to the causal organization. The proposed test generalizes the initial/final nodes consistency test of causal graph-based diagnostic schemes when causal relations are quantitatively and dynamically described and when variables are not simply associated with deviation from their nominal value but with the dynamics of their associated residuals.

Note that X_1 and X_2 are not necessarily detection variables; the error on X_1 (or X_2) may not be significant enough — in terms of the selected detection criteria — to be detected, but may nevertheless be sufficient to induce the error on X_3. Thus, even if X_1 and X_2 are missed at the detection stage, they may be proposed as root variables by the isolation step, and this result makes diagnosis reasoning more robust.

The generic rules that generalise the example in Eq. (14) and provide the order-of-magnitude solution of Eqs. (12) and (13) are:

- $(\forall k \in P_i,\ \varepsilon_i^k \cong \varepsilon_i)$: primary fault incriminating X_i,
- $(\varepsilon_i^k \ll \varepsilon_i)$ and $(\forall j \in P_i,\ j \neq k,\ \varepsilon_i^j \cong \varepsilon_i)$: secondary fault incriminating X_k;
- $(\exists P_i^{\mathrm{r}} \subset P_i)$ such that $(\varepsilon_i^{P_i^{\mathrm{r}}} \ll \varepsilon_i)$ and $(\forall k \in P_i^{\mathrm{r}},\ \varepsilon_i^k \approx \varepsilon_i)$: secondary fault incriminating $X_k,\ k \in P_i^{\mathrm{r}}$. (15)

The order of magnitude relations used to interpret the consistency test as proposed in Eq. (15) for isolation illustrates the idea that symbolic reasoning on real numbers is often sufficient to solve a practical case. However, implementing simple rules as proposed in Eq. (15) requires a proper formalization. The mathematical treatment of the order of magnitude relations $\ll$, $\approx$ and $\cong$ is

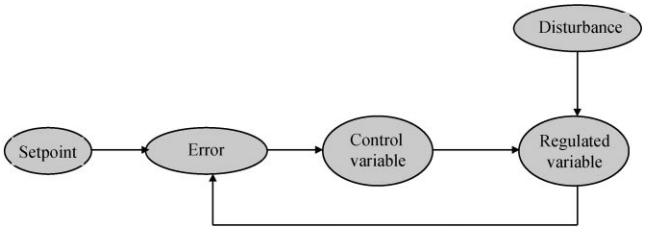


Fig. 4. Non-causal graph of a control loop.

achieved through a fuzzy set semantics and discussed in Section 4.

The test Eq. (15) is a first step in the isolation procedure; causality allows the local analysis of a variable X_i to be included in a recursive strategy for following the paths on the graph. The consistency test on X_i is followed by tests on its predecessors, and so on. The test is recursively carried out until an incriminated simulation error ε_j is no longer causally linked with at least one of its predecessors' errors; X_j is then declared root variable. The diagnostic result may be an input arc to the source variable X_j — corresponding to at least the partial disappearance of a plant function — or a disturbance directly affecting the source variable X_j; the malfunction may thus correspond either to a component failure or to a non measurable disturbance, a sensor or actuator fault. The other faulty detected nodes in the propagation graph are only variables influenced by the effects of the primary fault(s). The complete propagation graph of the fault effects can be plotted in this way, with the graph source(s) identified as the cause(s).

Finally, through the filtering procedure, a variable that has not been detected as faulty because of an inappropriately selected detection threshold can nevertheless be identified as a variable responsible for the fault, and can thus belong to the fault propagation sub-graph. This provides considerable flexibility in the choice of detection thresholds and ensures a very robust detection procedure (Montmain, 1992).

Causal modeling requires particular consideration of propagation of phenomena in loops. For material feedback loops, the propagation delay is consistent with the causal nature of the simulation. When the delay is nil, as in the case of a control loop (Fig. 4) the naturally associated graph is not causal.

Moreover, this graph would provide explanations of control loop behavior that are of no interest to the operator: it is not pertinent to explain the behavior of a regulated variable for which the setpoint has first been modified by a change in the control signal, modified again by the evolution of the error, and again by the evolution of the regulated variable. A representation that directly explains the behavior of a regulated variable, either by a change in the setpoint or by a disturbance is illustrated in Fig. 5a and is a causal representation (Leyval et al., 1994). Representations using causal graphs easily allow for changes in the structure of a process. Consider the same example of a control loop: the sub-graphs related to the behavior of the regulated variables must be constructed according to the state — closed or open — of the loop. Each control loop is therefore broken down into two sub-graphs, one corresponding to the closed-loop causal structure described above and the other to the open-loop structure (Fig. 5b) that is easily derived from the former.

Thanks to this causal transformation of a loop with a nil delay, fault isolation does not present any particular difficulty, and the previous consistency test may be applied between any disturbance and the control variable or the regulated variable, which retains an explanatory power. Moreover, the test Eq. (15) remains valid when used in closed loop (see Combastel, Gentil and Rognon (1999) for an application to a closed-loop electrical drive).

3.3. *Fault filtering use*

Once a propagation sub-graph has been identified, any subsequent faults will be tested using the same consistency test to determine whether they correspond to the occurrence of a new fault, or whether they are only the consequences of faults previously detected and accounted for; this is known as fault filtering or progressive monitoring. Propagation of the effects of the fault will lead to new simulation errors (and consequently to new

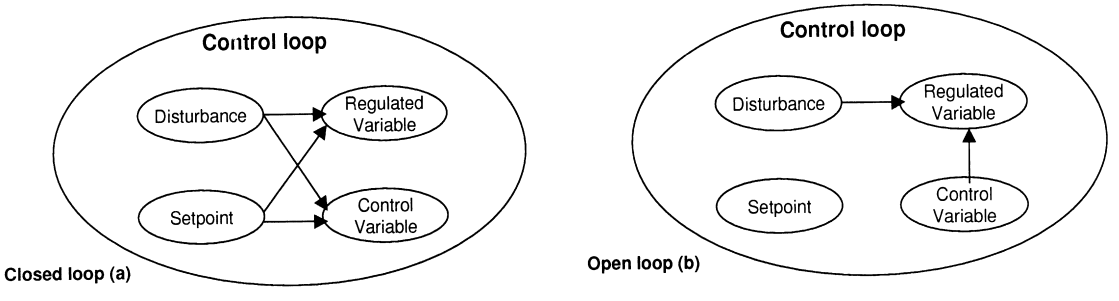


Fig. 5. Causal sub-graphs of closed (a) and open (b) control loops with a disturbance affecting the output.

detection variables), which will become the new terminal nodes on the fault propagation graph. With each measurement acquisition, the fault graph evolves as the effects of the fault propagate dynamically. With dynamic monitoring of the fault signature, the evolution of its consequences on the process can be explained continuously by the same source.

It is worth noting that the localization algorithm proposed in this paper does not consist simply in linking variables which have been detected in defect and which are related by an arc of the causal graph in order to declare that the root of this sub-graph is the source variable. It relies on a consistency test taking advantage of the dynamic properties of the arcs. It recursively tests every antecedent of a detected variable until a local fault has been proved, even if the antecedent is not presently detected faulty. In this way, a source fault that has been corrected by the operator but whose effects are still propagating due to long delays would still be considered as the explanation of the faulty variables.

Dynamic fault signature recognition is thus integrated in the operating tools in control rooms. Displaying the fault propagation sub-graph on the control interface constitutes an explanatory provision that is highly appreciated by operators (Evsukoff, Montmain & Gentil, 1998).

3.4. Extending the method

Consider again the general system in Eq. (1). Each equation in the system has the following form:

$$\dot{x}_i = g_i((\mathbf{x}_j)_{P_i}, x_i, \mathbf{u}_i), \tag{16}$$

where $\mathbf{u}_i$ represents the input vector directly affecting variable X_i and $(\mathbf{x}_j)_{P_i}$ the vector of the P_i direct antecedents of X_i. In normal operation, the process output verifies the following equation:

$$\dot{\bar{x}}_i = g_i((\bar{\mathbf{x}}_j)_{P_i}, \bar{x}_i, \mathbf{u}_i). \tag{17}$$

A predicted output on reconfiguration of the predecessor X_p is defined from Eq. (6)

$$\dot{x}_i^p = g_i((\mathbf{x}_j)P_{i/p}, \bar{x}_p, x_i^p, \mathbf{u}_i). \tag{18}$$

From Eqs. (4) and (7) after a first-order development of the $g_i((\bar{\mathbf{x}}_j)_{P_i}, \bar{x}_i, \mathbf{u}_i) - g_i((\mathbf{x}_j)_{P_i}, x_i, \mathbf{u}_i)$ and $g_i((\bar{\mathbf{x}}_j)_{P_i}, \bar{x}_i, \mathbf{u}_i) - g_i((\mathbf{x}_j)_{P_i/p}, \bar{x}_p, x_i^p, \mathbf{u}_i)$ terms, the simulation and prediction error dynamics are then written as

$$\dot{\varepsilon}_i = \frac{\partial g_i}{\partial x_i} \cdot \varepsilon_i + \sum_{j \in \boldsymbol{P}_i} \frac{\partial g_i}{\partial x_j} \cdot \varepsilon_j + \|h\| \cdot \varepsilon(h), \tag{19}$$

$$\dot{\varepsilon}_i^p = \frac{\partial g_i}{\partial x_i} \cdot \varepsilon_i^p + \sum_{j \in \boldsymbol{P}_i/p} \frac{\partial g_i}{\partial x_j} \cdot \varepsilon_j + \|h'\| \cdot \varepsilon(h'), \tag{20}$$

where $\|h\| \cdot \varepsilon(h)$ and $\|h'\| \cdot \varepsilon(h')$ are the residual terms of the development.

Integration yields the following equations describing the temporal behavior of the residuals:

$$\varepsilon_i(t) = \mathrm{e}^{-\varphi(t)}\left[\varepsilon_i(t0) + \sum_{j \in \boldsymbol{P}_i} \int_{t0}^{t} \mathrm{e}^{\varphi(\tau)} \frac{\partial g_i}{\partial x_j} \varepsilon_j(\tau)\,\mathrm{d}\tau + \int_{t0}^{t} \mathrm{e}^{\varphi(\tau)} \|h\| \varepsilon(h)\,\mathrm{d}\tau\right], \tag{21}$$

$$\varepsilon_i^p(t) = \mathrm{e}^{-\varphi(t)}\left[\varepsilon_i(t0) + \sum_{j \in \boldsymbol{P}_i/p} \int_{t0}^{t} \mathrm{e}^{\varphi(\tau)} \frac{\partial g_i}{\partial x_j} \varepsilon_j(\tau)\,\mathrm{d}\tau + \int_{t0}^{t} \mathrm{e}^{\varphi(\tau)} \|h'\| \varepsilon(h')\,\mathrm{d}\tau\right], \tag{22}$$

where $\varphi(t) = -\int_{t0}^{t} (\partial g_i / \partial x_i)\,\mathrm{d}\tau$. The first-order difference gives

$$\varepsilon_i(t) - \varepsilon_i^p(t) = \mathrm{e}^{-\varphi(t)} \int_{t0}^{t} \mathrm{e}^{\varphi(\tau)} \frac{\partial g_i}{\partial x_p} \varepsilon_k(\tau)\,\mathrm{d}\tau. \tag{23}$$

As in the particular case of a linear simulator, it is thus possible to obtain the contributions of each predecessor to the relevant simulation error. The principle proposed here to generate the residuals is not related to a particular representation or to a particular solution method; it remains valid not only for event driven simulation (Montmain et al., 1994a; Montmain & Leyval, 1994b), but also for difference equations (Evsukoff, Montmain & Gentil, 1997) and for continuous nonlinear differential equations (Vadam, Montmain & Cassar, 1997).

In practice, it is only necessary to calculate the open-loop simulation error Eq. (4) and the prediction errors Eqs. (7) and (9) obtained simply by successive substitution of the actual measured process value for each variable in the simulator equation, in order to determine whether it is a primary fault or a simple consequence (secondary fault). The algorithm is then applied recursively to the predecessors considered responsible for the identified error. Because of the approximations resulting from the developments used — among other reasons —, the layer of approximate reasoning necessary for a pertinent interpretation of the errors involved — as proposed in Eq. (15) and developed in Section 4 — is particularly warranted in the case of nonlinear systems.

4. Approximate reasoning for interpreting residuals and isolating faults

This section describes the mathematical formalism used to model the order of magnitude relations in the decision-making rules in Eq. (15). It is explained why and how this test can be modeled as a decision-making process in imprecise context.

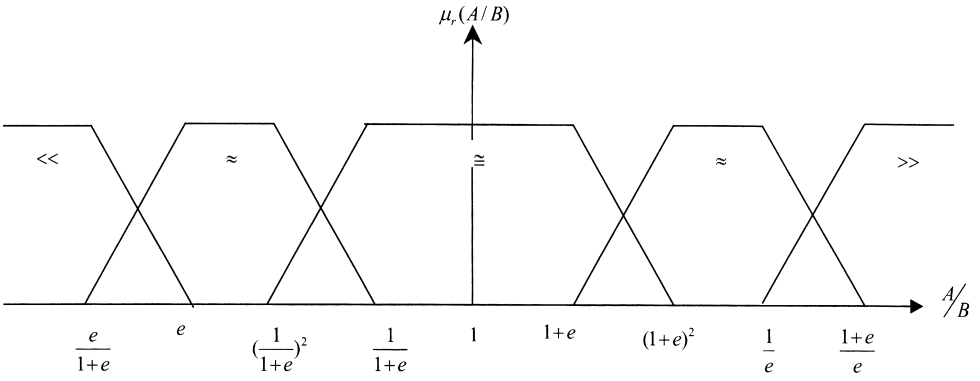


Fig. 6. Fuzzy representation and order of magnitude.

The consistency test fundamentally relies on Eqs. (4), (7), (9) and rules Eq. (15). Solving these equations in a purely numerical way would be the most accurate way to process the initial/final nodes consistency test if the model was perfect and the measurements were very precise. To solve this problem, thresholds defining the zero value of the residuals are generally introduced in numerical methods. Nevertheless the purely numerical solving cannot be carried out correctly if the thresholds of these various residuals, which are connected by arcs, are poorly balanced. When a malfunction or a nonmeasurable disturbance occurs, inappropriate thresholds may prevent the defect from being detected immediately on the root variable, or may upset the chronological sequence of fault occurrences in the causal chain: the primary fault may be detected first on one of its successors on the graph. To avoid a wrong diagnosis, these thresholds should be determined by using the very accurate quantitative information obtained by sensitivity analysis. This analysis is a rather dissuasive prospect for a process with a large number of variables. In fact the zero value is a vague concept and should be modeled as such.

Instead of associating an element of $\{0,1\}$ with a residual, it thus seems reasonable to use a real number in [0,1], which may be interpreted as the degree of possibility that the equation associated with the residual is in fact violated, considering the context. Fuzzy sets constitute a simple tool for creating the interface between low-level numerical data and high-level symbolic knowledge, inasmuch as they account for the continuous nature of the variables manipulated in a symbolic representation.

Moreover, Dubois (1983) indicated that the incentive to using fuzzy sets was the need to represent non-stringent specifications such as flexible constraints (for which slight violations are permissible). In any event, the idea is to achieve robust diagnosis by avoiding the sudden discontinuities that would arise if precise limits were introduced for the sets constituting the specification: sudden transition from permissible to unacceptable values, from values for which a procedure is applicable to nearly identical values for which it is no longer applicable.

As a conclusion, finding the fault propagation path must therefore be considered as an analysis in an imprecise context due to inappropriate thresholds or more generally imprecise measurements and modeling approximations. The idea beyond this remark is that like in many design or diagnostic activities in process engineering, order-of-magnitude reasoning is not only often sufficient in solving practical cases but also more realistic and safer. Thus the practical order of magnitude solving of Eq. (15) corresponds to this idea that symbolic reasoning on real numbers is often used by human beings.

In this imprecise decision-making environment, the $\ll$, $\approx$ and $\cong$ relations must firstly be interpreted in vague terms. Mavrovouniotis and Stephanopoulos (1988) proposed seven primitive binary order-of-magnitude relations. An "$A\,r\,B$" relation is equivalent to "$(A/B)\,r\,1$" and can be modeled as a fuzzy interval for the (A/B) ratio. Interpreting the relations in Eq. (15), only three order-of-magnitude relations are defined, using a single parameter e (Fig. 6):

- ' $\cong$ ': nearly identical to
- ' $\approx$ ': of the same order of magnitude
- $\ll$: negligible.

Threshold e has an upper limit of 0.4656 set by the constraint $1/e > (1+e)^2$. Otherwise, the value of e could be assumed to be 0.1, which corresponds to the common idea that an order of magnitude denotes roughly a factor of 10. Following this idea, Eq. (15) is replaced by

- $(\forall k \in P_i, |\varepsilon_i^k|/|\varepsilon_i| \cong 1)$: primary fault incriminating X_i;
- $(|\varepsilon_i^k|/|\varepsilon_i| \ll 1)$ and $(\forall j \in P_i, j \neq k, |\varepsilon_i^j|/|\varepsilon_i| \cong 1)$: secondary fault incriminating X_k; (24)
- $(\exists P_i^r \subset P_i)$ such that $(|\varepsilon_i^{P_i^r}|/|\varepsilon_i| \ll 1)$ and $(\forall k \in P_i^r, |\varepsilon_i^k|/\varepsilon_i \approx 1)$: secondary faults incriminating X_k, where $k \in P_i^r$.

It is now possible to give the formal model of the decision-making process that consists in determining whether an antecedent is responsible for the fault on one of its successors.

Fuzzy logic allows mathematical modeling of decision-making for imprecise and uncertain conditions, i.e. to assist the decision-maker in selecting an action to produce the desired consequences with respect to the assigned (possibly flexible) criteria and (possibly vague) constraints.

In a known environment, each action $a \in A$ (where A is the set of possible actions) has a known consequence described by a series of values $[m_1(a), m_2(a), \ldots, m_p(a)]$; $m_i(a)$ measures action a in the sense of criterion i; m_i defines an application of A in an objective scale X_i, and the set of consequences X is identified with the Cartesian product $X_1 \times X_2 \times \cdots \times X_p$. The decision-maker's objective for criterion i is a fuzzy set G_i of X_i such that $\forall x_i \in X_i$, $\mu_{G_i}(x_i)$ is the degree of compatibility between the decision-maker's objective and the value x_i describing the consequence (Dubois, 1983). Given the objective G_i and the criterion $i(m_i)$, each action a may be assessed for its expediency with regard to the objective G_i by the membership function μ_{γ_i} such that

$$\mu_{\gamma_i}(a) = \mu_{G_i}(m_i(a)). \tag{25}$$

When X is identified with $X_1 \times X_2 \times \cdots \times X_p$, the set of acceptable decisions D is defined as a fuzzy subset of the set of possible actions A, obtained by aggregating the sets of best actions γ_i based on the partial objectives $G_{i=1,p}$ constructed by Eq. (25). The membership function μ_D is thus such that

$$\begin{aligned} \forall a, \quad \mu_D(a) &= h[\mu_{\gamma_1}(a), \ldots, \mu_{\gamma_p}(a)] \\ &= h[\mu_{G_1}(m_1(a)), \ldots, \mu_{G_p}(m_p(a)], \end{aligned} \tag{26}$$

where h is a fuzzy set operator connective to be determined (Dubois & Prade, 1985).

Here, for a given X_i and its associated simulation error ε_i, an action a consists in *suspecting an antecedent* X_k, $k \in P_i^r$, of being responsible for ε_i; A is the set of possible X_k *Suspicions* and D is the associated fuzzy subset of *reasonable suspicions*; the measures m_i are identified with the simulation and prediction errors, and the partial objectives G_i are the order-of-magnitude relations constituting Eq. (24). Thus, for example, the membership function $\mu_{|\varepsilon_i^k|/|\varepsilon_i| \ll 1}(\varepsilon_i^k, \varepsilon_i)$ corresponds to the degree of relevance of the symbolic fact (partial clue in the suspicion examination) "ε_i^k *is negligible compared to* ε_i" to the situation described by the numerical values of ε_i^k and ε_i. Finally, μ_D (X_k *Suspicion*) is the value of the membership function to the *reasonable suspects* fuzzy set D for the element X_k and can be seen as the *suspicion degree* of X_k in the X_i simulation error responsibility. X_k is finally *suspected* when its suspicion degree is beyond a given threshold.

Considering the decision model proposed in Eq. (26), decision-making rules incriminating X_k in Eq. (24) thus become for each $k \in P_i$:

$$\mu_{\text{Reasonable Suspicions}}(X_k \text{suspicion}) = h\begin{bmatrix} \mu_{|\varepsilon_i^k|/|\varepsilon_i| \ll 1}(\varepsilon_i^k, \varepsilon_i), \{\mu_{|\varepsilon_i^j|/|\varepsilon_i| \cong 1}(\varepsilon_i^j, \varepsilon_i)\}_{j \in P_i, j \neq k}, \\ \mu_{|\varepsilon_i^k|/|\varepsilon_i| \approx 1}(\varepsilon_i^k, \varepsilon_i), \{\mu_{|\varepsilon_i^j|/|\varepsilon_i| \approx 1}(\varepsilon_i^j, \varepsilon_i)\}_{j \in P_i^r, j \neq k}, \mu_{|\varepsilon_i^{P_i^r}|/|\varepsilon_i| \ll 1}(\varepsilon_i^{P_i^r}, \varepsilon_i) \end{bmatrix} \tag{27}$$

Necessary conditions on operator h are enumerated in Appendix B. Three main decision-making attitudes may be modeled using the aggregation function: conjunction, compromise and disjunction, although all possible intermediate attitudes may be imagined. Identification techniques of h operator and application to fault isolation modeled as a decision-making process are proposed in Montmain and Gentil (1996). The choice of h is made explicit in Eq. (28), where '$\wedge$' indicates a conjunction and '$\vee$' a disjunction.

$$\begin{aligned} &\mu_{\text{Reasonable Suspicions}}(X_k \text{suspicion}) \\ &= \left[\bigwedge_{j \in P_i, j \neq k} \{\mu_{|\varepsilon_i^j|/|\varepsilon_i| \cong 1}(\varepsilon_i^j, \varepsilon_i)\} \wedge \mu_{|\varepsilon_i^k|/|\varepsilon_i| \ll 1}(\varepsilon_i^k, \varepsilon_i) \right] \\ &\vee \left[\bigwedge_{j \in P_i^r, j \neq k} \{\mu_{|\varepsilon_i^j|/|\varepsilon_i| \approx 1}(\varepsilon_i^j, \varepsilon_i)\} \wedge \mu_{|\varepsilon_i^k|/|\varepsilon_i| \approx 1}(\varepsilon_i^k, \varepsilon_i) \right. \\ &\quad \left. \wedge \mu_{|\varepsilon_i^{P_i^r}|/|\varepsilon_i| \ll 1}(\varepsilon_i^{P_i^r}, \varepsilon_i) \right]. \end{aligned} \tag{28}$$

The fuzzy conjunction and disjunction operators were selected as follows: $u \wedge v = u \cdot v$ and $u \vee v = u + v - u \cdot v$. The use of *probabilistic* operators for aggregation can be justified from a purely mathematical standpoint: their strict monotonic character can notably be a useful property in this application. Their other interest lies in their quick processing for on line implementation.

Finally the algorithm could be summed up as follows:

- simulation errors ε_i are calculated at each acquisition time through Eq. (4); they provide the corresponding detection variable set;
- for each detection variable X_i, prediction errors after reconfiguration are established through Eqs. (7) and (9);
  - for each antecedent of X_i all the membership functions associated to the order-of-magnitude relations in Eq. (28) are evaluated;
  - the degree of suspicion of each antecedent of X_i is evaluated by Eq. (28);

- the test is recursively carried out with the suspected antecedents as new detection variables;
- the fault propagation sub-graph is achieved and upgraded at each acquisition time.

5. Description of an application

The methodology described in the preceding sections was implemented in a nuclear fuel reprocessing plant application. Spent fuel is reprocessed to recover the uranium and plutonium still present in the fuel rods irradiated in nuclear reactors and to isolate the residual fission products.

The process considered in this application consists of two-pulsed columns. A pulsed column is a liquid–liquid extractor. The aqueous input stream comprising two flows, $Q0500$ (spent fuel and nitric acid) and $QG100$, is mixed with the organic solution $Q1010$ in the column. The acid solution outflow stream $QE120$ contains only the fission products. The organic solution outflow stream $QG600$ containing uranium and plutonium is extracted by overflow. The mixture is subjected to periodic pressure pulses $PRE801$. The equilibrium between the various components is regulated by the level $NIRE$.

The two-pulsed columns are coupled top-to-bottom so the organic outflow from the extraction column becomes the organic inflow $QG600$ to the scrubbing column. The aqueous outflow $QG100$ from the scrubbing column is recycled to the extraction column feed port (Fig. 7). A corresponding simplified causal graph is proposed in Fig. 8, which involves only 32 variables. The complete application addresses the detection and isolation of about 55 variables (flow rates, pressures, column weights and levels) and five local regulation loops (mainly levels and flow rates). This gives an idea of the applicability of the method to a complex process.

A precise simulator of the plant is available for operator training. Several types of faults are managed by the simulator: sensor or actuator faults (drift, noise, offset), operating faults (e.g. a poorly selected setpoint) or process faults (plugging, fouling, inadequate head). About 60 faults can be simulated. Multiple fault scenarios can be solved without any major difficulty (excepted compensatory effects of some faults). The results discussed below were obtained not on the actual process, but with this simulator. The training simulator, which is a reference in the domain, provides a highly realistic description of process behavior, corresponding to what is actually observed by the operators on the control interfaces.

This training simulator plays the role of the process — $\{\bar{x}_i\}$ — in the study presented in this paper. The simulated behaviors — $\{x_i\}$ — are obtained with a causal dynamic simulator where the arcs between a variable X_i and its predecessors X_j are identified as difference equations as proposed in Eq. (3).

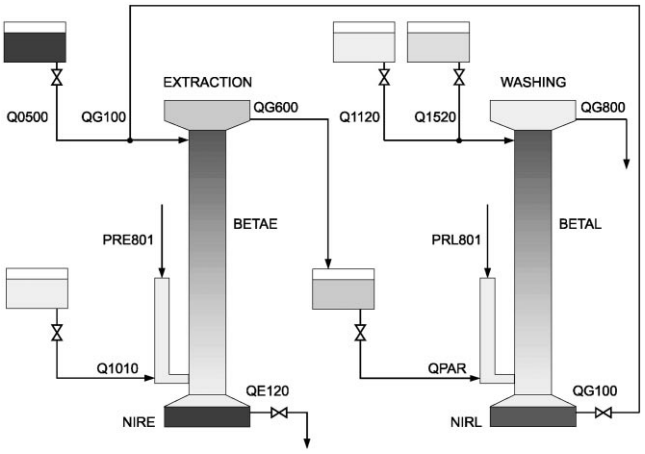


Fig. 7. Schematic diagram of plant.

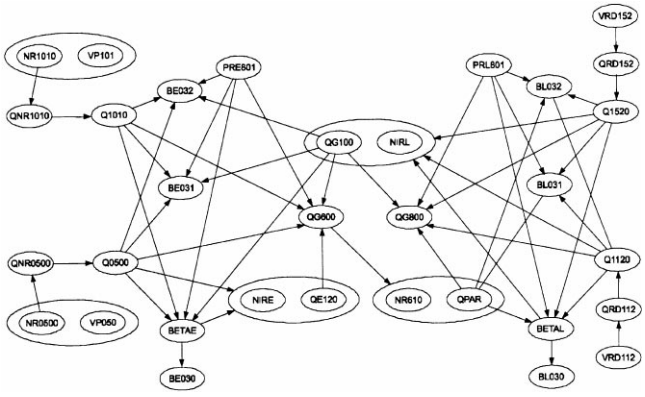
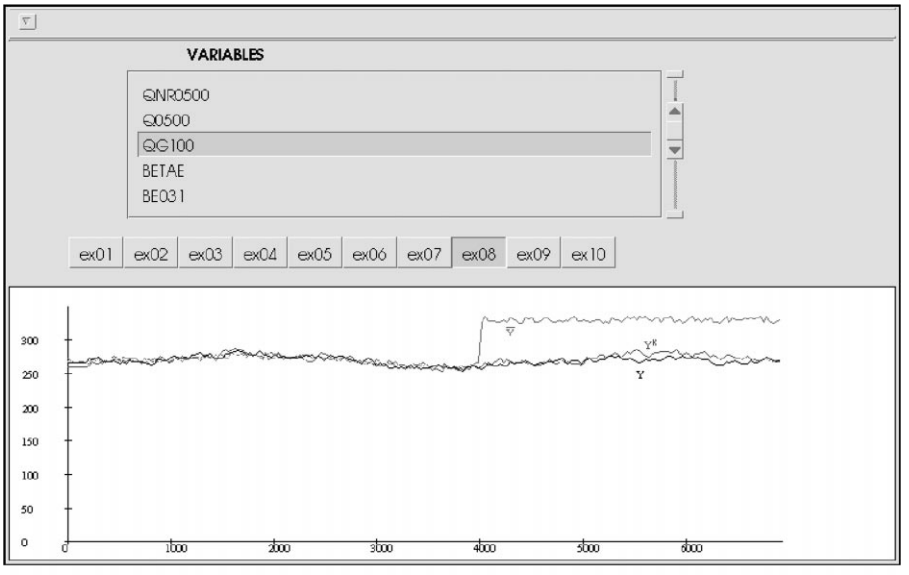
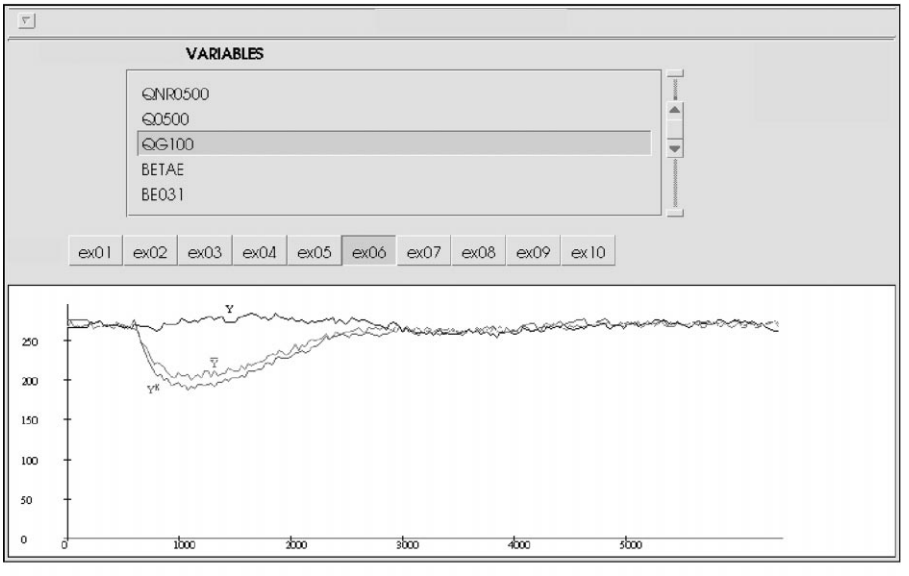


Fig. 8. Part of the associated hydraulic causal graph.

Fig. 9 shows two examples of consistency tests in the case of a fault on flow rate $QG100$: a local fault (a) and a fault due to a pressure fault (b). The training simulator output $QG100$, the measurement $\overline{QG100}$ and the causal graph prediction $QG100^k$ clearly illustrate the difference between the two situations. In the case of the local fault — sensor fault on the aqueous outflow $QG100$ from the scrubbing column — whatever the reconfigured antecedent $k \in \{BETAL, Q1120, Q1520\}$ of $QG100$, it can be verified (Fig. 9a) that $\varepsilon^k_{QG100} \cong \varepsilon_{QG100}$ ($\bar{y} - y^k \cong \bar{y} - y$ in the figure) and the local fault is concluded Eq. (28). In the second case (Fig. 9b) the illustrated example corresponds to a pressure pulse primary fault of the scrubbing column ($PRL801$). As a quasi immediate effect, the heavy phase inner flow of the scrubbing column is affected, and consequently the hold-up $BETAL$ and the column weights ($BL030$, $BL031$ and $BL032$) characteristic of the emulsion operating mode are abnormally modified. To keep the level ($NIRL$) of the settler at the bottom of the column constant, the heavy phase outlet ($QG100$) is finally subject



(a)



(b)

Fig. 9. Isolation from variable $QG100$. (a) Local fault, $QG100 \neq \overline{QG100}$ and $QG100 \cong QG100^k$. (b) Upstream fault, $QG100 \neq \overline{QG100}$ and $\overline{QG100} \cong QG100^k$.

to abnormal variations. When $k \in \{Q1120, Q1520\}$ it could be shown that $\varepsilon^k_{QG100} \cong \varepsilon_{QG100}$ and $Q1120$ and $Q1520$ are not incriminated. When k is $BETAL$ (Fig. 9b). $\varepsilon^{BETAL}_{QG100} \ll \varepsilon_{QG100}$ ($\bar{y} - y^k \ll \bar{y} - y$ in the figure) and $BETAL$ is incriminated. Then the consistency test is recursively carried out until the root $PRL801$ is found. In fact, the $QG100$ abnormal deviation is a very late observation of the initial fault on the pressure. It can still be noticed that $BETAL$ is difficult to evaluate accurately and the corresponding threshold on ε_{BETAL} cannot be as strict as the ones on ε_{PRL801} (predecessor) or ε_{QG100} (successor). This is a typical case when there is a poor balance between the thresholds of linked variables. However, the use of fuzzy set theory in the order of magnitude reasoning avoids a wrong diagnosis and provides the right root variable even when $BETAL$ has not been declared as faulty by Eq. (5) in the detection step. The fault on $QG100$ will finally affect the extraction column and a later view of the fault propagation path is showed in Fig. 10 (Evsukoff et al., 1998).

Let us now consider a drift fault on $Q0500$ — it corresponds to a slow danaïd clogging. This fault occurs during a ramp set point change, which is rather difficult to detect before a steady state has been reached. The first observed effect is an abnormal deviation on $Q0500$. Instantaneously the organic phase outflow $QG600$ is

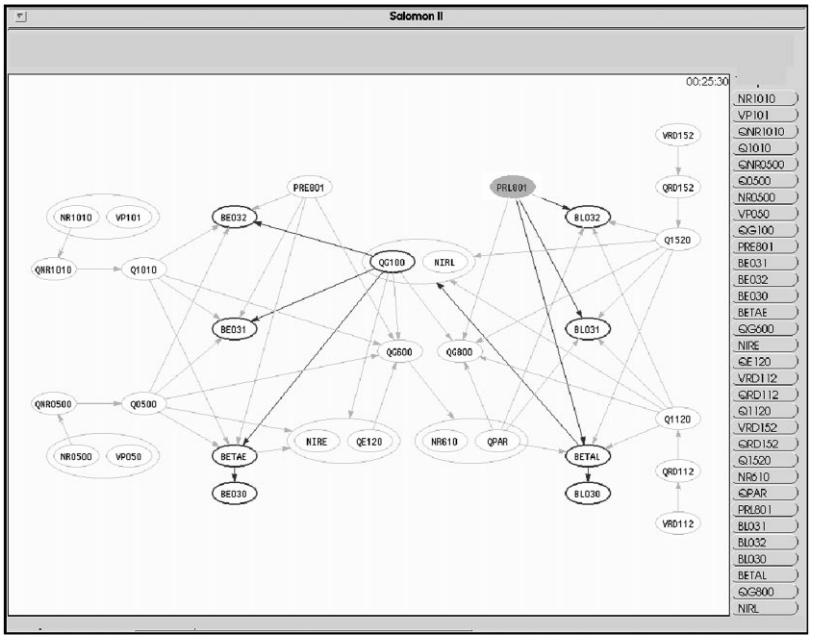


Fig. 10. View of the fault propagation path.

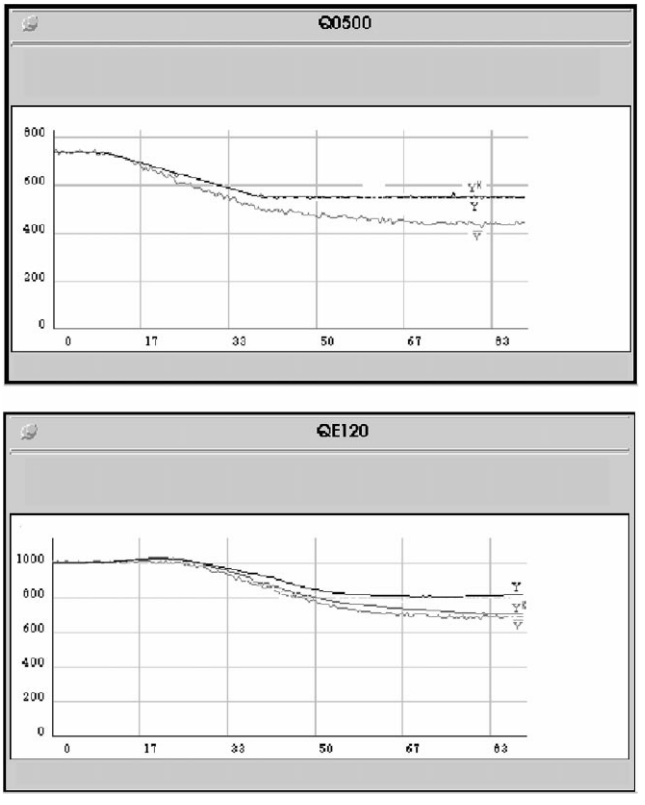


Fig. 11. Consistency tests on variables QO500 and QE120.

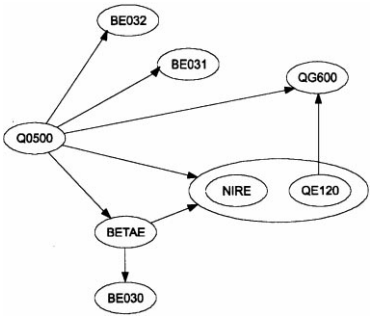


Fig. 12. Fault propagation path.

modified due to the inflow–outflow balance. The system verifies that the faults on $Q0500$ and $QG600$ are linked: the fault on $QG600$ is only a secondary effect. It is a cascading fault case and not a multiple fault scheme.

Then the hold-up $BETAE$ and the column weights ($BE030, BE031$ and $BE032$) characteristic of the extraction emulsion are involved with the hydraulic dynamics of the column. When the aqueous inner flow disturbance reaches the settler at the bottom of the column, the aqueous outflow $QE120$ that regulates the settler level $NIRE$ is affected. Fig. 11 illustrates this step. It is shown that $\varepsilon_{Q0500}^{QNR0500} \cong \varepsilon_{Q0500}$ and consequently $Q0500$ is a primary fault; moreover the test on $QE120$ concludes that $\varepsilon_{QE120}^{Q0500} \approx \varepsilon_{QE120}$, $\varepsilon_{QE120}^{BETAE} \approx \varepsilon_{QE120}$ and $\varepsilon_{QE120}^{\{Q0500,\, BETAE\}} \ll \varepsilon_{QE120}$: $Q0500$ and $BETAE$ are both incriminated. Finally, the last test between $Q0500$ and $BETAE$ gives the fault propagation path in Fig. 12. The dynamic monitoring of the primary fault effects is necessary to ensure a continuous assessment of the disturbed functions that allows the operators to revise their initial decision if necessary and to undertake the appropriate reaction at any time — i.e. from any observed fault signature.

The dynamic analysis for fault isolation and fault filtering is successfully achieved and the order of magnitude reasoning relying on rules Eqs. (24) and (28) is proved to be sufficient to interpret the generated structured residuals. It corresponds not only to the idea that human beings often efficiently reason symbolically on numerical data, but as explained in Section 4, this qualitative management of quantitative equations eliminates two major problems:

- the illusive and irrelevant precision of purely quantitative methods as soon as models are imprecise and measurements uncertain;
- the poor detectability and distinguishability capacities of purely qualitative approaches as well as the difficult question of time management.

The process presented is this section is a continuous one with slow dynamics (time responses are of the order of 30 min) and an acquisition time of 30 s. Despite the recursive graph search, the suspect sub-graph elaboration never exceeds 1 s in this application where the number of nodes is 55 and the maximal number of antecedents is 5. This gives a rough idea of the algorithm complexity.

6. Conclusion

The method proposed in this paper uses the precision of numerical techniques while benefiting from causal knowledge of the process to implement diagnostic reasoning suitable for industrial processes. The isolation power of numeric methods was obtained here without algebraic manipulations to obtain a structured residual set. Causal knowledge is used to apply the classic concepts of model-based diagnosis — simulation, prediction and measurement coherence test — to local sub-models; isolation is then obtained naturally by testing the coherence of the local models with the measured values. Recursivity makes it possible to test numerous variables with very simple models and produces a chain of suspected variables. Fault filtering may be considered as a supplementary means of increasing the robustness of the detection phase.

The method can be used to construct a complete fault detection, isolation and filtering system for less constraining representations than a state representation. When a simulator implementing differential equations is already available, it is not even necessary to know the analytical structure of the model to perform the isolation: having established the causal structure of the simulator, with only a few software properties — e.g. provision for switching the differential equation inputs — it may be considered simply as a black box with switchable inputs. In this configuration, the method makes fault isolation possible without extending or further refining the existing models of the industrial facility.

When a numerical simulator is not available a model may be directly constructed as a causal model. A top-down analysis from the balance equations is used to add nodes and arcs until the graph includes sufficient details. Approximate temporal parameters may be used to quantify the arcs, or classic identification procedures may be used to assign difference equations to them. The causal graph is not necessarily easy to develop, and requires a detailed physical analysis, but this in turn is the source of its explanatory capability. Moreover, as the process itself evolves, the causal graph is modified without FDI algorithm re-coding.

In this diagnostic methodology for online supervision of dynamic complex processes, the fault filtering function constitutes an essential process supervision support intended more for the control operator than for an *upgraded* SCADA able to reconfigure the plant in any malfunctioning case. The progressive monitoring of observed fault signatures is considered to be as essential as an optimal FDI scheme with theoretical fault signatures.

Algorithms based on the causality principle are simple, dissociated from the model, and avoid all the algebraic manipulations that confuse the diagnostic reasoning process. The fact that a numeric control is based on a mathematical model that is meaningless to the operator is not a serious defect inasmuch as it concerns a closed loop in the process. For computerized supervision, however, the operator is part of the decision-making loop; the mathematical artifices used by the supervision system must therefore be given a cognitive transcription corresponding to the operator's representation and reasoning in faulty situations. The method proposed here is dedicated to supervision and diagnosis: it uses a causal process breakdown to minimize the required calculations and to provide the diagnostic system with the explanatory character that is crucial in ensuring satisfactory perception in the control room. The information supplied by the system is explained in a manner pertinent to the knowledge domains of the operating staff.

A fault detection and isolation system in a complex plant is primarily an operating support. The causal model-based process supervision concept is consistent with this point of view. Displaying the behavior of the model and process in a historical log provides a basic tool for understanding the diagnosis. Displaying the graph and the fault propagation paths is another tool that has been highly appreciated by experienced operators in a nuclear process.

Appendix A: Glossary

x_i	simulated behavior of variable X_i
u_i	input vector directly affecting variable X_i

f_i	fault directly affecting variable X_i
$\bar{x}_i$	actual behavior of X_i
x_i^p	predicted evolution of X_i when antecedent X_p has been reconfigured: x_p has been replaced by its measured evolution $\bar{x}_p$ in the calculus of x_i
P_i	set of all subscripts j of the predecessors X_j of X_i
P_i^r	set of all subscripts j of the predecessors X_j of X_i that have been reconfigured
U_i	set of subscripts j of the process inputs directly affecting X_i
ε_i	simulation error or open-loop residual for variable X_i
ε_i^p	prediction error on reconfiguration of the predecessor X_p
$\varepsilon_i^{P_i^r}$	prediction error on reconfiguration of the predecessors of variable X_i associated with P_i^r
$\cong$	nearly identical
$\approx$	of the same order of magnitude
$\ll$	negligible
$\mu_F(\varepsilon)$	degree of relevance of symbolic fact F to the situation described by the numeric value ε
h	fuzzy set aggregation connective

Appendix B: Properties of fuzzy set aggregation connectives

Necessary conditions on operator h are the following: h is continuous;

$$h(0,0,\ldots,0) = 0 \quad \text{and} \quad h(1,1,\ldots,1) = 1; \tag{B.1}$$

$$\forall (u_i, v_i) \in [0,1]^2, \text{ if } u_i \geq v_i \text{ then } h(u_1,\ldots,u_p) \geq h(v_1,\ldots,v_p).$$

Three main decision-making attitudes may be modeled using the aggregation function: conjunction, compromise and disjunction, although all possible intermediate attitudes may be imagined.

For an operator h expressing that all the criteria are met simultaneously, a natural axiom is:

$$\forall (u_1,\ldots,u_p), \quad h(u_1,\ldots,u_p) \leq \min(u_1,\ldots,u_p), \tag{B.2}$$

i.e. the overall evaluation of an action cannot be better than the worst of the partial evaluations. These operators are conjunctions. The main associative conjunctions are

$$\min(u_1,\ldots,u_p).$$

$$\prod_{i=1}^{p} u_i,$$

$$\max\left(0, \sum_{i=1}^{p} u_i - p + 1\right). \tag{B.3}$$

To express the redundancy of the objectives, operator h must meet the following condition:

$$\forall (u_1,\ldots,u_p), \ \max(u_1,\ldots,u_p) \leq h(u_1,\ldots,u_p), \tag{B.4}$$

i.e. the overall evaluation is determined by the best of the partial evaluations. These operators are disjunctions. The most frequently used are

$$\max(u_1,\ldots,u_p)$$

$$1 - \prod_{i=1}^{p} (1 - u_i),$$

$$\min\left(1, \sum_{i=1}^{p} u_i\right). \tag{B.5}$$

Operator h is a compromise when the following axiom is verified:

$$\forall h(u_1,\ldots,u_p), \ \min(u_1,\ldots,u_p) \leq h(u_1,\ldots,u_p) \leq \max(u_1,\ldots,u_p). \tag{B.6}$$

All conjunctive operators can be covered by the Yager operator family

$$Y_q(u_i) = 1 - \min\left(1, \left(\sum_{i=1}^{p} (1-u_i)^q\right)^{1/q}\right) \quad \text{for } q \geq 0. \tag{B.7}$$

For example,

$$\lim_{q \to \infty} Y_q(u_i) = \min(u_i),$$

$$Y_1(u_i) = \max\left(0, \sum_{i=1}^{p} u_i - p + 1\right).$$

Parameters are assigned to the disjunctive operators by the associated conorm family

$$CY_q(u_i) = \min\left(1, \left(\sum_{i=1}^{p} u_i^q\right)^{1/q}\right) \quad \text{for } q > 0. \tag{B.8}$$

The compromise operators are described as follows:

$$Ym_q(u_i) = \left(\frac{\sum_{i=1}^{p} u_i^q}{p}\right)^{1/q}. \tag{B.9}$$

Notable q values include the arithmetic mean $Ym_1(u_i)$ and the harmonic mean $Ym_{-1}(u_i)$; when $q \to 0$, Eq. (B.9) is the geometric mean.

The symmetrical characteristic of h does not imply that the aggregation is symmetrical. Indeed, aggregation relies on the fuzzy sets describing the partial objectives, and dissymmetry may easily be introduced among criteria by manipulating the membership functions. The above aggregation functions can be generalized with balancing coefficients in order to easily introduce natural dissymmetries.

References

Chen, J., Patton, R. J., & Zhang, H. -Y. (1996). Design of unknown input observers and robust fault detection filters. *International Journal of Control*, *63*(1), 85–105.

Combastel, C., Gentil, S., & Rognon, J. -P. (1999). Fault detection and isolation using local models, comparison with unknown input observers. *ECC 99*, Karlsruhe, Germany.

Davis, R. (1983). Diagnosis via causal reasoning: Paths of interaction and the locality principle. *American Association for Artificial Intelligence Conference* (pp. 88–94).

Dubois, D. (1983). *Modèles mathématiques de l'imprécis et de l'incertain en vue d'applications aux techniques d'aide à la décision*. Ph.D. thesis, Institut National Polytechnique de Grenoble.

Dubois, D., & Prade, H. (1985). A review of fuzzy set aggregation connectives. *Information Sciences*, *36*, 85–121.

Evsukoff, A., Montmain, J., & Gentil, S. (1997). Dynamic model and causal knowledge-based fault detection and isolation. *IFAC safeprocess '97* (pp. 699–704), Hull, UK.

Evsukoff, A., Montmain, J., & Gentil, S. (1998). Causal model based supervising and training. *IFAC workshop on line fault detection and supervision in the chemical industries*, Lyon, France.

Frank, P. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy, a survey and some new results. *Automatica*, *26*(3), 459–474.

Frank, P. (1991). Fault diagnosis in dynamic system using software redundancy. *Revue européenne Diagnostic et Sûreté de fonctionnement, Hermès*, *1*(2), 113–143.

Frank, P. (1994). Application of fuzzy logic to process supervision and fault diagnosis. *Safeprocess'94*, *IFAC symposium on fault detection*, *Supervision and safety for technical processes*, vol. 2 (pp. 531–537), Espoo, Finland.

Frank, P. (1996). Analytical and qualitative model-based fault diagnosis, a survey and some new results. *European Journal of Control*, *1*(2), 6–28.

Gertler, J. (1997). Fault detection and isolation using parity relations. *Control Engineering Practice*, *5*(5), 653–661.

Gertler, J. (1998). *Fault detection and diagnosis in engineering systems*. New York, USA: Marcel Dekker.

Iri, M., Aoki, K., O'Shima, E., & Matsuyama, H. (1980). Graphical approach to the problem of locating the origin of the system failure. *Journal of Operations Research Society of Japan*, *23*(4), 295–312.

Isermann, R. (1993). Fault diagnosis of machines via parameter estimation and knowledge processing-Tutorial paper. *Automatica*, *29*(4), 815–835.

Isermann, R., & Ballé, P. (1997). Trends in the application of model-based fault detection and diagnosis of technical processes. *Control Engineering Practice*, *5*(5), 709–719.

Leyval, L., & Ledoux, A. (1991). Simulation for a nuclear retreatment facility. *IFAC safeprocess' 91*, Baden-Baden, Germany.

Leyval, L., Gentil, S., & Feray-Beaumont, S. (1994). Model based causal reasoning for process supervision. *Automatica*, *30*(8), 1295–1306.

Mavrovouniotis, M., & Stephanopoulos, G. (1988). Formal order-of-magnitude reasoning in process engineering. *Computer and Chemical Engineering*, *12*(9/10), 867–880.

Montmain, J. (1992). *Interprétation qualitative de simulations pour le diagnostic en ligne de procédés continus*. Ph.D. thesis, Institut National Polytechnique de Grenoble.

Montmain, J. (1997). From Diapason research program to its industrial application in nuclear fuel reprocessing. *IFAC safeprocess '97* (pp. 209–216), Hull, UK.

Montmain, J. (1998). Operators aids: Automation and supervision. *Ninth symposium on information control in manufacturing*, *INCOM'98* (pp. 215–221), Nancy-Metz, France.

Montmain, J., & Gentil, S. (1993). Interprétation qualitative pour le diagnostic en ligne. *Revue Européenne Diagnostic et Sûreté de Fonctionnement*, *3*(1), 23–45.

Montmain, J., Leyval, L., & Gentil, S. (1994a). Qualitative analysis for decision making in supervision of industrial continuous processes. *Mathematics and Computers in Simulation*, *36*, 149–163.

Montmain, J., & Leyval, L. (1994b). Causal graphs for model based diagnosis, *IFAC safeprocess'94*, Espoo (Finland).

Montmain, J., & Gentil, S. (1996). Operation support for alarm filtering. *IEEE Conference CESA'96*, *Computational Engineering in Systems Applications*, Lille, France.

Mosterman, P., Biswas, G., & Narasimham, S. (1997). Measurement selection and diagnosability of complex physical systems. *Eight international workshop on principles of diagnosis*, *DX'97* (pp. 79–86), Mont Saint Michel, France.

Mosterman, P., Biswas, G., & Manders, E. (1998). A comprehensive framework for model-based diagnosis. *Ninth international workshop on principles of diagnosis*, *DX'98* (pp. 86–93), Cape Code, USA.

MQ&D coordinated by P. Dague (1995). Qualitative reasoning: A survey of techniques and applications, *AI Communications the European Journal of AI*, *8* (3/4), 119–192.

Palowitch, B., & Kramer, M. (1986). The application of a knowledge based expert system to chemical plant fault diagnosis. *American control conference* (pp. 646–651).

Patton, R., & Chen, J. (1991). A review of parity space approaches to fault diagnosis. *Safeprocess '91*, *IFAC symposium on fault detection*, *supervision and safety for technical processes*, vol. 1 (pp. 239–255), Baden-Baden, Germany.

Penalva, J. M., Coudouneau, L., Leyval, L., & Montmain, J. (1993). DIAPASON: A supervision support system. *IEEE Expert Intelligent Systems and their Applications*, *8*(5), 57–65.

Rasmussen, J. (1993). Diagnostic reasoning in action. *IEEE Transactions on Systems Man and Cybernetics*, *23*(4), 981–991.

Shiozaki, J., Matsuyama, H., Tano, K., & O'Shima, E. (1985). Fault diagnosis of chemical processes by the use of signed, directed graphs: Extension to five-range patterns of abnormality. *International Chemical Engineering*, *25*(4), 651–659.

Vadam, O., Montmain, J., & Cassar, J.-P. (1997). Fault detection using parallel simulations. *IFAC safeprocess '97* (pp. 121–126), Hull, UK.

Yu, C., & Lee, C. (1991). Fault diagnosis based on qualitative/quantitative process knowledge. *A.I.Ch.E. Journal*, *37*(4), 617–627.