FCC DIAGNOSIS USING SEVERAL CAUSAL AND KNOWLEDGE BASED MODELS

B. Heim[1], S. Gentil[2], B. Celse[1], S. Cauvin[3], L. Travé-Massuyès[4]

[1]IFP - BP 3 - 69390 Vernaison, France
[2]LAG (CNRS – INPG – UJF)- BP46 - 38402 Saint Martin d'Hères Cedex 1, FRANCE
[3]IFP- 1 & 4, avenue de Bois Préau - 92852 Rueil Malmaison, FRANCE
[4] LAAS, 7 avenue du colonel Roche, 31000 Toulouse, FRANCE
email: Bruno.Heim@ifp.fr, Sylviane.Gentil@inpg.fr, Benoit.Celse@ifp.fr, Sylvie.Cauvin@ifp.fr, Louise@laas.fr

Abstract: In order to deal with the complexity of the diagnosis of FCC pilot plants, several modelling approaches were developed, combined and tested on-line. Two causal modelling approaches were investigated based on control loop analysis and on the detailed equations describing the behaviour of the process. These models are used on-line to detect faults on process variables. Information on the components of the system allows faults on physical components to be isolated. Then using expert knowledge, information is given to the operator. This paper details the different kinds of models, their use in the diagnosis module and a case study on the IFP FCC pilot plant. This work is conducted as a part of the Chem project[1].

Keywords: Fault detection, fault isolation, fault identification, model based diagnosis, causal graphs, knowledge based system.

1. INTRODUCTION

The FCC (Fluid Catalytic Cracking) refinery process contains a physical loop generated by the catalyst circulation in the different components. This makes the understanding of the spreading of faults very complex. Furthermore, the FCC pilot plant is representative of the increasing complexity of highly automated modern plants (89 sensors and 53 regulators). In spite of this complexity, a model-based diagnosis approach was chosen because it ensures completeness of the diagnosis and means that a physical explanation of the diagnosis can be given to the operators (Isermann and Ballé, 1997).

A causal model is particularly appropriated to complex processes because it is an explanatory tool for supporting fault detection and isolation. Moreover, it can be easily modified.
During normal behaviour, a causal model describes qualitatively and quantitatively the influences among process variables. A possible representation of a causal model is a causal graph made of nodes and directed arcs. Nodes represent variables and arcs represent influences among variables. The information carried by the arcs can be purely qualitative (signs of influence for instance) but it can also be quantitative: gains for a static representation or transfer functions to take time into consideration (Leyval, *et al.*, 1994) (Travé-Massuyès and Milne 1996).
In this paper, a quantitative transfer function is associated with the arcs. Given a variable x that influences a variable y, value for y can be generated either based on a model value for x (simulation) or based on a measured value for x (prediction). Values are propagated from node to node easily.

[1] CHEM: "Advanced Decision Support System for chemical/petrochemical processes" Project is funded by the European Community under the Competitive and Sustainable Growth programme of the Fifth RTD Framework Programme (1998-2002) under contract G1RD-CT-2001-00466. See www.cordis.lu or www.chem-dss.org

Section one presents two causal modelling approaches. The first one is based on classical first principle models and produces a deep model. The equations need to be rearranged to obtain a causal ordering (Iwasaki and Simon, 1986). The second one focuses on control loops analysis. Representation of control loops and generation of generic transfer functions are considered.
Section 2 details the diagnosis module. It can be divided into three sub-modules for fault detection, isolation and identification (Isermann and Ballé, 1997).
First, symptoms are generated by comparing the causal model predicted outputs with the measured values. (Montmain and Gentil, 2000).
Secondly, the isolation module determines a list of suspected physical components. Each arc of the causal model is connected to a set of components. This qualitative knowledge is used by a hitting set algorithm to isolate the source component (Cordier *and al*, 2000).
Finally, the knowledge based module, generates operator messages. Each physical component is associated with qualitative models of abnormal behaviour containing human operator knowledge about this component. When a component is suspected to be abnormal then its models of abnormal behaviour are analysed to generate more information to the operator.
Section 3 describes the application of this methodology to a part of the FCC pilot plant. The advantages and drawbacks of a deep causal model and an expert causal model are compared and a complete scenario is described.

2. CAUSAL MODELLING

2.1 *Description of modelling approaches*

As described in (Heim *and al*, 2002) the following steps are required to generate a deep causal model:

1) Identification of the physical system with respect to its environment (exogenous phenomena),
2) Division into sub-systems,
3) Assignment of a configuration to each sub-system,
4) Identification of the set V of variables required to describe the system,
5) Identification of the set R of physical relations among those variables,
6) Connection of each relation to physical components,
7) Determination of causality (application of a causal ordering algorithm),
8) Reduction (elimination of non measured variable),
9) Approximation (elimination of negligible phenomena),
10) Quantification (identification of transfer function parameters).

When dealing with the deep approach, a classical first principle model is used at step 5. Therefore equations are explicitly known.
The bipartite graph $G=(V \cup R, A)$ is defined, where A is the set of influences. The causal ordering (step 7) arises from determining a perfect matching in G (Travé-Massuyès and Pons 1997).
On the opposite, the expert approach represents the system from a control point of view, starting with the description of the main control loops and adding incrementally the representation of secondary control loops or disturbances.
Fig. 2 shows the causal graph related to the control loop described in Fig. 1. Set-point SP acts on the manipulated variable MV via the controller R. MV and disturbances D act on the regulated variables RV via process H.

Fig. 1. Regulation loop

The major drawback of this representation is the double link between RV and MV, that creates the indefinite looping of the influences. A way to break this loop is to use the causal representation shown in Fig. 2. This causal graph explains the modification in RV or MV due to modifications of SP or D.

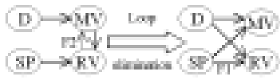

Fig. 2. Loop elimination

This representation can be extended to more complicated cascaded control loops (Heim *and al*, 2000).
Transfer functions in this last representation can be computed easily with some reasonable hypotheses. F_1 is the closed loop transfer function between SP and RV and F_2 is the open loop transfer function of the process between MV and RV. They are assumed to be known by the process expert and have the simple form described in (1).This approximation is justified in the application described in §3.

$$F(s) = \frac{g.e^{-d.s}}{(1+T.s)^n} \tag{1}$$

where parameters are a gain g, a delay d, a time constant T and the order n.
The influence of D on RV is described by the generic transfer (2). The transfer functions between SP or D and MV differ only by their sign and are represented by the generic transfer (3).

$$FD(s) = \frac{g'.s.e^{-d'.s}}{(1+T.s)^n} \tag{2}$$

$$FMV(s) = \frac{1+\tau s}{(1+Ts)^n} = \frac{F_1(s)}{F_2(s)} \tag{3}$$

2.2 Deep and expert approach comparison

Steps (1) to (5) and step (10) are common to the expert and to the deep approach. In the expert approach causality is automatically deduced from causal loop representation (Fig. 2). In the deep approach steps (7) to (9) are necessary to determine causality and simplify the model. Step (6) is common to the two approaches and is necessary to isolate faults on physical components (cf. 3.2)

- In the deep causal model, the physical equations that rule the system must be known. In the expert causal model only behavioural relations are known. As the deep causal model contains more knowledge than the expert causal model, the diagnosis is therefore more complete.
- The expert causal model can be obtained rapidly describing only influences of set-points on manipulated and regulated variables. It can be refined adding influences of disturbances on these variables, then steps (7) to (9) become necessary. The deep causal model requires a long modelling time.
- To quantify the expert causal model, closed and open loop experiments have to be carried out in order to obtain the parameters of transfer functions F1 and F2. F1 can also be obtained directly from F2 when the controller R is known.
- To quantify the deep causal model, closed and open loop experiments also have to be carried out. As parameters often appear in more than one equation, the identification task is less time consuming than in the expert approach.
- The deep causal graph ensures the completeness of the list of components associated with influences. Elementary relationships are used to describe the process, therefore the set of components associated with an arc is minimised. With the expert causal graph approach the completeness of the diagnosis is assured only when all the disturbances on the variables are known.
- Relationships used in the expert causal model are not elementary. Therefore the list of components associated with the influences is not minimised. For instance the influence of SP on RV is associated to the controller R (c.f. Fig. 1), the process P and the sensor RV. In the deep causal model, description is based on elementary relationships, therefore R appears in one relation and H in another relation. Distinction between a fault on R and a fault on H is then possible.

3. DIAGNOSTIC MODULE

3.1 Fault detection and isolation on process variables

The causal model is used for fault detection on a process variable. Let $Y_i(t)$ be the measured value of each node Y_i of the causal graph. The causal model provides a simulated reference $Y_i(t)^*$ and a predicted reference $\hat{Y}_i(t)$. The simulated reference represents the value of a node computed from the set-points and the exogenous disturbances acting on the process. The predicted reference represents the value of a node computed from the measured values of the direct antecedent nodes. For example, in Fig. 5, the simulated reference of F1 is obtained from SP and SF1 and the predicted reference of F1 is obtained from P and OPV1. With these two values, two residuals are defined by comparison with the variable measurement:

$$r_{Yi}(t) = \left| Y_i(t) - Y_i^*(t) \right| \tag{4}$$

$$\lambda_{Yi}(t) = \left| Y_i(t) - \hat{Y}_i(t) \right| \tag{5}$$

The global residual $r_{Yi}(t)$ and the local residual $\lambda_{Yi}(t)$ are used for fault detection and isolation on process variables. The causal diagnostic methodology consists in deciding for each node if the fault is local (and thus explains all the other observed discrepancies) or if the fault is upstream (and is thus explained by the fault on another antecedent variable).

This is done by testing the coherency between the measurements and the model with a logical reasoning. The results of a Boolean reasoning on these residuals are shown in Table 1. In this table, “1” indicates that the value of the residual is greater than a threshold and “0” that the residual is smaller than this threshold.

Table 1 Boolean reasoning on residuals

$r_{Yi}(t)$	$\lambda_{Yi}(t)$	Fault
1	1	Local
1	0	Upstream
0	1	Impossible
0	0	No Fault

Let $r_{n,Yi}(t)$ and $\lambda_{n,Yi}(t)$ be the normalised global and local residuals respectively given by (6) and (7).

$$r_{n,Yi}(t) = \min \left| \frac{r_{Yi}(t)}{a_{Yi}}, 1 \right| \tag{6}$$

$$\lambda_{n,Yi}(t) = \min \left| \frac{\lambda_{Yi}(t)}{b_{Yi}}, 1 \right| \tag{7}$$

Threshold values on residuals a_{Yi} and b_{Yi}, can be deduced from alarm thresholds on Y_i. Parameter a_{Yi} is the highest accepted value for the global residual and parameter b_{Yi} is the highest accepted value for the local residual. The gradual evolution of $r_{n,Yi}(t)$ and $\lambda_{n,Yi}(t)$ from 0 to 1 characterises the evolution of the variable from a normal state to an undesirable state. This transition is translated in terms of a colour grade in the user interface. $r_{n,Yi}(t)$ is used to colour the contour of the nodes: the contour of the node is red when $r_{n,Yi}(t) = 1$ and green when $r_{n,Yi}(t) = 0$. The arcs influencing a variable are red if $\lambda_{n,Yi}(t) = 1$ and green if $\lambda_{n,Yi}(t) = 0$. In between, transition colours are used to visualise residual evolutions.

3.2 Fault isolation on physical components

Each relation of the causal graph is associated with one or several physical component(s). Associating influences with physical components enables the fault on physical components to be isolated. A component of the physical system is associated with a relationship, if and only if, this relationship determines its (or part of its) behaviour. For instance, the influences P->F2 and OPV2->F2 (see equation 8 and Fig. 5) are both associated with the valve V2. Therefore, if V2 has a normal behaviour, then F2 must be consistent with P and OPV2. On the other hand, if the residual λ_{F2} is high, then V2 is suspected to have an abnormal behaviour. When a local fault is detected on a variable, then the set of components associated with the arcs between this variable and all the parents of this variable determine a conflict. A conflict is a set of components, at least one of which behaves abnormally. Hitting set reference algorithm is used to isolate faults on components. (Minimal) diagnoses can be generated from (minimal) conflicts using a hitting sets algorithm. A diagnosis is hence a set of components such that its intersection with all the conflict sets is not empty. In this paper the assumption is made that a fault always manifests itself.

3.3 Knowledge based fault identification

When a physical component is suspected, the aim of this module is to refine the diagnosis about this component using a knowledge base about its abnormal behaviour. To complete this task, each physical component is associated with models of abnormal behaviour obtained from human operator knowledge. This model takes the form of an event-tree model (Fig. 3). Given a component suspected by the previous isolation module 3.2), signals that constitute symptoms, characterising faults, on this component are analysed.
Symptoms and faults are symbolised by objects. The relationship between the origin of the fault and the symptom are symbolised by directed arcs between objects, from symptoms to faults (Fig. 3). Parameters qualify symptoms (variable identity, amplitude, frequency ...). A fault can be determined by different symptoms, therefore symptoms are linked to faults using logical tests (AND, OR).
A message is sent to the operator interface when diagnosis can be refined using the expert knowledge.
The process history, from records of maintenance and repairs, constitutes a source of heuristic information. Each fault is characterised by numerical symptoms or observed qualitative symptoms. Numerical symptoms are used to determine automatically possible fault origin(s). Tendencies (increases , decreases), shapes (pulse & , oscillation , steps &) and comparison (high >, and low <) are detected through signal analysis.
If there are negligible and complicated phenomena that are not modelled in the causal model, model of abnormal behaviour can be used to refine diagnosis of those phenomena. Consequently variables that are analysed in the third module are those represented in the causal model but also other measured process variables.
Qualitative symptoms (odour, smell, colour, noises, vibration ...) can be provided in the messages to the operator interface and can be analysed by the human operator to confirm the diagnosis physically.

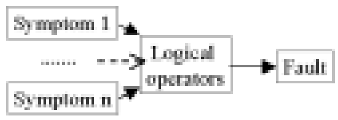

Fig. 3. Qualitative abnormal model structure

4. APPLICATION TO A PART OF THE FCC PILOT PLANT

The diagnostic system was implemented with the real time G2 software by Gensym. G2 employs an object oriented methodology that makes it easy to modify a causal graph by creating new arcs and nodes.
This section presents the application of the two approaches (deep and expert) to a part of an FCC pilot plant. A real data based scenario of abnormal behaviour is finally presented.
An FCC process is a refinery process which receives multiple feeds consisting of high boiling point components from several other refinery process units. It cracks these streams into lighter components. It is composed of many subsystems (two regenerators, a reactor, a separation column, pipes, valves ...). The catalyst circulates in a physical closed loop: it goes from the stripper to the 1st regenerator then to the 2nd regenerator and finally comes back to the stripper. The feed is put in contact with the catalyst, and immediately, catalyst and reaction products fall into the stripper. The FCC pilot plant which is about 15 meters high has exactly the same components as a real FCC process.
A causal model of the whole FCC pilot plant was developed. The methodologies presented above have been tested on real data from an FCC pilot process and are at present being tested on-line.
These methodologies are applied on the sub-system illustrated by Fig. 4. It is composed of the tank T_1 (first regenerator) and two valves V_1 and V_2. T_1 contains essentially nitrogen at pressure P and at temperature T. Nitrogen input flow is F_1 and output flow is F_2. A controller RV_2 acts on V_2 to maintain P at its set-point SP. A regulator RV_1 acts on V_1 to maintain F_1 at its set-point SF_1.

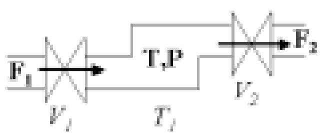

Fig. 4. Regenerator sub-system

4.1 Deep approach

Physical relations that rule the studied system are given by explicit equations (8) to (12). Only the pressure balance is considered. R_1 and R_2 respectively represent the transfer functions of regulators RV1 and RV2. Disturbances T, PNNT and P_{atm} are not measured and will thus not be considered in the causal graphs presented later: they are eliminated from the causal model at step 9 in § 2.1. NNT is the nitrogen network. With:
F_1,F_2 moles/s, P pressure (Pa), T temperature (K), P_{atm} atmospheric pressure (Pa), P_{NNT} nitrogen network pressure (Pa), Ts sampling time (s), V volume of T_1 (m^3), $OPV_{i=1,2}$ is the opening value for V_i (%), f and g non linear relationships.

$$\frac{dP}{dt} = \frac{T}{V} * T_S * (F_1 - f(OPV_2)\sqrt{(P^2 - P_{atm}{}^2)} \quad (8)$$

$$F_1 = f(OPV_1)\sqrt{(P^2 - P_{NNT}{}^2)} \quad (9)$$

$$F_2 = g(OPV_2)\sqrt{(P^2 - P_{atm}{}^2)} \quad (10)$$

$$OPV_1 = R_1(SF_1 - F_1) \quad (11)$$

$$OPV_2 = R_2(SP - P) \quad (12)$$

The application of the deep approach methodology to equations (8) to (12) provides the causal graph in Fig. 5. Each arc of the causal graph is associated to a (some) physical component(s). This association, illustrated in Fig. 5, is used for diagnostic purposes (cf. § 3.2). In Fig. 5 association of influences with sensors is not mentioned. To ensure the finest diagnosis, each arc must also be associated with the sensors involved (for instance F1->P is associated with both sensors F1 and P).

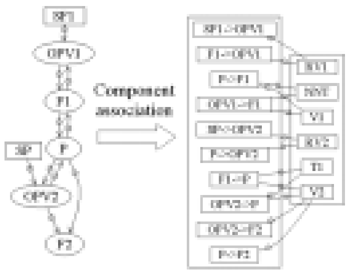

Fig. 5. Deep causal graph

4.2 Expert approach

The expert approach is simple. First it allows a systematic and simplified representation of the control loop (cf. Fig. 2). Second it is not necessary to describe explicitly the physical equations that rules the process. Indeed it is sufficient either to ask the process expert which variables are disturbances of the manipulated value (and thus of the regulated value) or ask the expert for formal relations that rules the process. In this section the influences of the set-point will first be considered. Then the influences of the disturbance will be dealt with.
First set-point SF1 influences both F1 and OPV1 and set-point SP influences both P and OPV2. Parameters of the transfer functions SF1->F1 and SP->P (relation named *F1* in § 2.1) can be easily obtained by performing closed loop steps on SF1 and SP. Performing open loop steps on OPV1 and OPV2 gives the relation F2 between OPV1 and F1 and between OPV2 and P. Then parameters of transfer functions SP1->OPV1 and SP->OPV2 (F1/F2) are obtained. At this step of the expert approach the numerical influences of the set-points are known and the causal graph in Fig. 6 can be used to generate variable values.
It is now necessary to identify which variables are disturbances for F1, OPV1, P and OPV2. If no physical knowledge is available about the process, then it is impossible to identify those disturbances. If physical relations (8) to (12) are formally known or if a process expert is able to determine which variables are disturbances, then the expert causal graph in Fig. 6 is obtained.

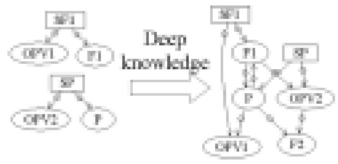

Fig. 6. Expert causal graph

Arcs {P->F2; OPV2->F2} are associated with the same components than those mentioned in Fig. 5. Arcs {SF1->F1; SF1->OPV1; P->F1; P->OPV1} and {SP->OPV2; SP->P; F1->P; F1->P} are respectively associated with {T1; V1; RV1, NNT} and with {T1; V2; RV2}.

4.3 Scenario description

An expert causal model was first tested rapidly for the FCC pilot plant fault diagnosis. It was sufficient to isolate faults on certain components. It gave encouraging results but was however limited because of its simplified process representation. To provide a more precise diagnosis, a deep causal model was built. The following scenario is thus based on the deep causal model.
The pipe between T1 and valve V1 that controls P is blocked. In Figure 5, we can see that faults are detected on OPV2, OPV1, F1, F2 and P. Faults are isolated on F2 and P.
Fig. 8 presents the whole implemented FCC pilot plant causal graph at the end of the scenario. Faults are detected on the grey variables in Fig. 8. Bold arcs influence variables on which faults are isolated.

- Arcs that influence F2 are associated with {V2} and with {P sensor, value OPV2}, therefore, the set {V2, P sensor, value OPV2} determines a first conflict. Arcs that influence P are associated with {T1, V2} and with {P sensor, F1 sensor, FOR1 sensor, value OPV2}. That is the second conflict. In this scenario, P sensor is not considered to be abnormal because the local residual of DPC in Fig. 8 is zero. Minimal diagnoses are thus {V2} and {value of OPV2}.
- Historical faults known by the operator about V2 are: an internal leakage of V2, an external leakage of

V2 and a blockage of V2. The only historical fault known about the value of OPV2 is the Intensity/Pressure converter of V2 being out of order. Abnormal behaviour model associated with a blockage of V2 is illustrated in Fig. 7. At the beginning of the scenario, P and OPV2 increases and at the end of the scenario, P and OPV2 are high.

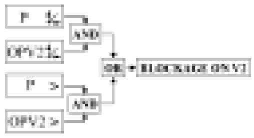

Fig. 7. Qualitative model of abnormal behaviour

This scenario occurred twice in real time. The beginning of the scenario cannot be exactly known. It is impossible to determine when V2 began to block. The only information that is available is measured values P and F2. Here the beginning of the scenario is arbitrarily chosen when the $\lambda_{n,P}(t)=1$, cf.(7). During the scenario, colours of nodes P and F2 slowly evolve from green to red for 40 minutes long. Without the diagnostic module, the fault was not isolated until the security components automatically stopped the process. After 5 minutes of abnormal behaviour, the diagnosis module isolated the fault. This illustrates the necessity of such modules.

Initially, the third module was tested on line, alone on the FCC pilot plant. The expert knowledge about each component was analysed simultaneously and many messages were sent to the operator to inform him of the findings. Focusing this module only on the components that are suspected to be faulty by the isolation module makes its use much more relevant. An average of three fault-trees for each component were built.

5. CONCLUSION

This paper presents a diagnostic module for an FCC pilot plant. Two causal modelling approaches were investigated and compared: a deep approach based on a precise description of the process with a first principle model and an expert approach based on control loop analysis. When using the expert approach, the diagnostic module is weaker but it can be obtained very quickly. The diagnostic module based on three complementary techniques was detailed. First, the fault detection module, based on a causal model of normal behaviour generates symptoms for each measured variable. Then, the isolation module generates a list of suspected physical components. Finally, the third module refines the diagnosis using historical qualitative operator knowledge of abnormal behaviour.

This diagnostic module was tested on 15 scenari based on real data and is at present being tested on line. It gives very encouraging results providing early and accurate detection. Within the scope of Chem, it is planned to integrate other techniques in the system in order to build a complete and integrated supervision application.

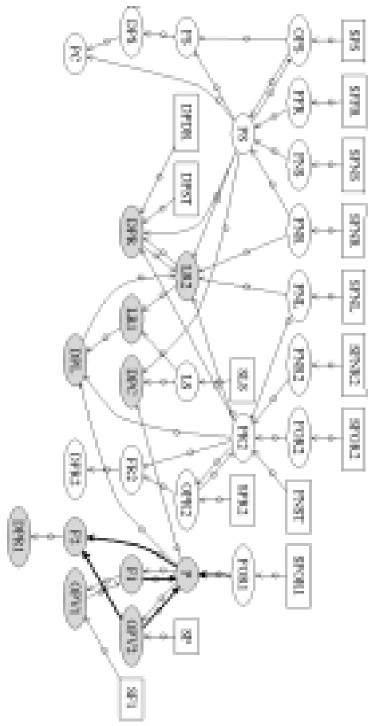

Fig. 8. FCC deep causal graph

REFERENCES

Cordier, M., P. Dague, M. Dumas, F. Levy, J. Montmain, M. Staroswiecki, L. Travé-Massuyès (2000). A comparative analysis of AI and control theory approaches to model-based diagnosis, *ECAI'00*, Berlin, Germany.

Heim, B., S. Cauvin, S. Gentil (2000). Causal and fuzzy reasoning methodology for cascaded loops diagnosis. *2000 IAR ICD Workshop devoted to Intelligent Control and Diagnosis,* Nancy, France.

Heim, B., S. Gentil, S. Cauvin, L. Travé-Massuyès, B. Braunschweig (2002). Fault diagnosis of a chemical process using causal uncertain model. *Prestigious Applications of Artificial Intelligence, subconference of the 15th European Conference on Artificial Intelligence,* Lyon, France.

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

Iwasaki, S. and H. Simon (1986). Causality in device behavior, *Artificial intelligence*, 29(1-3):3-32.

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

Montmain, J. and S. Gentil (2000). Dynamic causal model diagnostic reasoning for on-line technical process supervision, *Automatica* 36, 1137-1152.

Travé-Massuyès, L. and R. Milne (1996). Diagnosis of dynamic systems based on explicit and implicit behavioural models : An application to gas turbines in esprit project tiger, *Applied Artificial Intelligence Journal*, 10(3).

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