ABNORMAL EVENTS MANAGEMENT IN COMPLEX PROCESS PLANTS: CHALLENGES AND OPPORTUNITIES IN INTELLIGENT SUPERVISORY CONTROL

Venkat Venkatasubramanian

Laboratory for Intelligent Process Systems,
School of Chemical Engineering,
Purdue University
West Lafayette, IN 47907, USA.

Abstract: One of the most important challenges facing control system engineers is the design and implementation of next-generation control systems that can assist operators in making supervisory control decisions such as in abnormal events management (AEM), start up and shut down, controller performance assessment and so on. Operator failure to exercise the appropriate supervisory control decisions often have an adverse effect on product quality, process safety, occupational health and environmental impact. The economic impact of such abnormal situations is enormous; an estimated $20 billion/year in losses in the petrochemical industries alone in the US. Furthermore, process safety, occupational health and environmental hazards are ever increasing in importance in response to heightening public concern and the resultant tightening of regulations. Thus, there exist considerable incentives in developing intelligent control systems that can provide automated operator assistance for supervisory control situations for complex process plants. People in the process industries view this as the next major challenge in control systems research, design and application. Since fault detection and diagnosis is an important first step in AEM, I start with an overview of the various approaches to fault diagnosis, before discussing the challenges and the encouraging emerging trends. Recent progress in this area has promising implications on the use of intelligent systems for inherently safer design, operator training, abnormal events management, and process hazards analysis.

Keywords: Fault Diagnosis, Abnormal Events Management, Process Hazards Analysis

1. INTRODUCTION

The discipline of process control has made tremendous advances, both in theory and practice, in the last three decades with the advent of computer control of complex processes. Low-level control actions, called *regulatory control*, which used to be performed by human operators are now routinely performed in an automated manner with the aid of computers with considerable success. With progress in distributed control and model predictive control systems, the benefits to various industrial segments such as chemical, petrochemical, cement, steel, power and desalination industries have been enormous. However, a very important control task in managing process plants still remains largely a manual activity, performed by human operators. This is the domain of *supervisory control*, where operators have to make quick decisions based on complex causal reasoning about the control of abnormal process situations, start up and shut down of processes/systems, optimal control strategies and so on.

However, this complete reliance on human operators to cope with such abnormal events and emergencies has become increasingly difficult due to several factors. It is difficult due to the broad scope of the diagnostic activity that encompasses a variety of malfunctions such as process unit failures, process unit degradation, parameter drifts and so on. It is further complicated by the size and complexity of modern process plants. For example, in a large process plant there may be as many as 1500 process variables observed every few seconds (Bailey, 1984) leading to information overload. In addition, often the emphasis is on quick diagnosis which poses certain constraints and demands on the diagnostic activity. Furthermore, the task of fault diagnosis is made difficult by the fact that the process measurements may often be insufficient, incomplete and/or unreliable due to a variety of causes such as sensor biases or failures.

Given such difficult conditions, it should come as no surprise that human operators tend to make erroneous decisions and take actions which make matters even worse, as reported in the literature. Industrial statistics show that about 70% of the industrial accidents are caused by human errors. These abnormal events have significant economic, safety and environmental impact. Despite advances in computer-based control of chemical plants, the fact that two of the worst ever chemical plant accidents, namely, Union Carbide's Bhopal, India, accident and Occidental Petroleum's Piper Alpha accident (Lees, 1996), happened in the 1980s is a troubling development.

Recent events have shown that such large-scale plant accidents are not just things of the past but continue to haunt us even today. In this regard, I would like to bring your attention to three important major incidents. The first was the explosion at the Kuwait Petrochemical's Mina Al-Ahmedhi refinery in June of 2000. Fortunately, the human casualties were relatively low but financially the incident ranks among the top ten worst accidents with an estimated damages of about $400 million. The next major recent incident was the explosion at the offshore oil platform of Petrobras, Brazil, and its subsequent sinking into the sea in March 2001. The estimated losses in this case are about $5 billion. And last, in Sept 2001, the AZF chemical plant, Toulouse, had a large explosion that killed dozens of people and hundreds were treated for various kinds of injuries. It also disrupted the operations of the local university and its chemical engineering department which were situated near the plant site. All these accidents were covered extensively by various news organizations including CNN, and hence the public awareness of the potential accidents in chemical plants and their consequences continues to remain heightened.

Further, industrial statistics show that even though major catastrophes and disasters from chemical plant failures may be infrequent, minor accidents are very common, occurring on a day to day basis, resulting in many occupational injuries, illnesses, and costing the society billions of dollars every year (Bureau of Labor Statistics, 1998; National Safety Council, 1999). It is estimated that the petrochemical industry in the U. S. incurs approximately $20 billion in losses due to poor abnormal events management (Nimmo,1995). The cost is much more when one includes similar situations in other industries such as pharmaceutical, specialty chemicals, power, desalination and so on. As a very recent example, Nucor Corporation Inc. is paying $100 million towards fines and remedies in a pollution control lawsuit. Similarly, accidents cost the British economy up to $27 billion every year (Laser, 2000).

All these concerns, and the lessons learnt from these accidents, have led the federal agencies in the U.S. to create tighter safety, health and environmental regulations. The Occupational Safety and Health Administration (OSHA) passed its PSM standard Title 29 CFR 1910.119, which requires all major chemical plant sites to perform process hazards analysis (PHA) (OSHA, 1992). In addition, EPA instituted the Risk Management Program (RMP) in 1995. Similar regulations are coming up in Europe as well (Laser, 2000). All these require the systematic identification of process hazards, their assessment and mitigation. Process Hazards Analysis is the systematic identification, evaluation and mitigation of potential process hazards which could endanger the health and safety of humans and cause serious economic losses. The importance of performing a comprehensive PHA is illustrated by Kletz (1988; 1991) with examples of industrial accidents that could have been prevented if only a thorough PHA had been performed earlier on that plant. AEM and PHA are two sides of the same coin. Both are concerned with process safety: AEM is concerned about diagnosing abnormal causal origins of adverse consequences while PHA deals with reasoning about adverse consequences from abnormal causes. Intelligent, real-time, operator support systems are seen as a way to address both AEM and PHA. The automation of process fault detection and diagnosis forms the first step in automating supervisory control and AEM.

Thus, here in lies the next grand challenge for control engineers. In the past, the control community showed how *regulatory control* can be automated using computers thereby removing it from the hands of human operators. This has led to great progress in product quality and consistency, process safety and process efficiency. The current grand challenge is in the automation of *supervisory control* using intelligent control systems, thereby providing human operators the assistance in this most pressing area of need. People in the process industries view this as the next major challenge in control systems research and application. There are, of course, a number of practical challenges in designing such systems due to several factors such as the complexity of process dynamics, lack of adequate models, incomplete and uncertain data, diverse sources of knowledge, amount of effort and expertise required to develop and maintain the systems etc. However, considerable progress has been made in all areas of this field. One is in the development of multiple state estimators technology that combines several dfferent approaches to process monitoring and diagnosis. The other is the progress made in the development of intelligent systems for PHA.

In this paper, I briefly review the various approaches for fault diagnosis. I will discuss a framework for comparing the different approaches to understand their relative strengths and weaknesses. I will also discuss the recent trends and speculate about future work. Due to the volume of literature on this subject and the limited length of this paper, this review is necessarily brief. For a more exhaustive treatment, the reader is referred to the review papers by Venkatasubramanian *et al.* (2002a; 2002b; 2002c). Recent progress in this area has promising implications on the use of intelligent systems for inherently safer design, operator training, abnormal situation management, process hazards analysis and optimal process operations. Intelligent control systems are poised to define the nature of process control research and application for the coming decade.

2. FAULT DIAGNOSTIC SYSTEMS: Some Preliminaries

The term *fault* is generally defined as a departure from an acceptable range of an observed variable or a calculated parameter associated with a process (Himmelblau, 1978). This defines a fault as a process abnormality or symptom, such as high temperature in a reactor or low product quality and so on. The underlying cause(s) of this abnormality, such as a failed coolant pump or a controller, is (are) called the *basic event*(*s*) or the *root cause*(*s*). The basic event is also referred to as a *malfunction* or a *failure*. Since one can view the task of diagnosis as a classification problem, the diagnostic system is also referred to as a diagnostic classifier. Figure 1 depicts the components of a general fault diagnostic framework. The figure shows a controlled process system and indicates the different sources of failures in it. In general, one has to deal with three classes of failures or malfunctions as described below:

Gross parameter changes in a model: Parameter changes arise when there is a disturbance entering the process from the environment through one or more exogenous variables. An example is the change in the heat transfer coefficient due to fouling of a heat exchanger.

Structural changes: Structural changes refer to changes in the model itself. They occur due to hard failures in equipment. An example is a controller failure which would imply that the manipulated variable is no longer functionally dependent on the controlled variable.

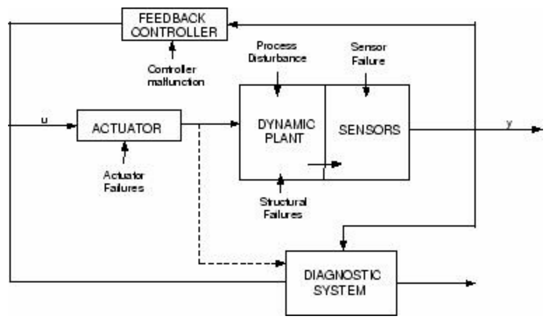

Fig. 1. A general diagnostic framework

Malfunctioning sensors and actuators: Gross errors usually occur with actuators and sensors. These could be due to a fixed failure, a constant bias (positive or negative) or an out-of-range failure.

Outside the scope of fault diagnosis are unstructured uncertainties, process noise and measurement noise. Unstructured uncertainties are mainly faults that are not modeled *a priori*. Process noise refers to the mismatch between the actual process and the predictions from model equations, whereas, measurement noise refers to high frequency additive component in the sensor measurements.

3. DESIRABLE FEATURES OF A FAULT DIAGNOSTIC SYSTEM

In order to compare various diagnostic approaches, it is useful to identify a set of desirable characteristics that a diagnostic system should possess. Then the different approaches may be evaluated against such a common set of requirements or standards. Though these characteristics will not usually be met by any single diagnostic method, they are useful to benchmark various methods in terms of the *a priori* information that is needed, reliability of solution, generality and efficiency of computation. In this context, one needs to understand two important concepts: completeness and resolution, before proceeding to the characteristics of a good diagnostic classifier. Whenever an abnormality occurs in a process, a general diagnostic classifier would come up with a set of hypotheses or faults that explains the abnormality. Completeness of a diagnostic classifier would require the actual fault(s) to be a subset of the proposed fault set. Resolution of a diagnostic classifier would require the fault set to be as minimal as possible.

(1) Quick detection and diagnosis:

The diagnostic system should respond quickly in detecting and diagnosing malfunctions. However, quick diagnosis and tolerable performance during normal operation are two conflicting goals (Willsky, 1976). A system that is designed to detect a failure (particularly abrupt changes) quickly will be sensitive to high frequency influences. This makes the system sensitive to noise and can lead to frequent false alarms during normal operation, which can be disruptive. This is analogous to the trade-off between robustness and performance noted in the control literature.

(2) Isolability:
Isolability is the ability of the diagnostic system to distinguish between different failures. Under ideal conditions free of noise and modeling uncertainties, this means that the diagnostic classifier should be able to generate output that is orthogonal to faults that have not occurred. There is also a trade-off between isolability and rejecting modeling uncertainties.

(3) Robustness:
One would like the diagnostic system to be robust to various noise and uncertainties. Its performance should degrade gracefully instead of failing totally and abruptly. This implies that the thresholds should be tuned conservatively. However, as noted earlier, this can affect performance.

(4) Novelty Identifiability:
One of the minimal requirements of a diagnostic system is to be able to decide, given current process conditions, whether the process is functioning normally or abnormally, and if abnormal, whether the cause is a known malfunction or an unknown, novel, malfunction. This criterion is known as novelty identifiability.

(5) Classification error estimate:
An important practical requirement for a diagnostic system is in building the user's confidence on its reliability. This could be greatly facilitated if the diagnostic system could provide *a priori* estimates on classification error that can occur. Such error measures would be useful to project confidence levels on the diagnostic decisions by the system giving the user a better feel for the reliability of the recommendations by the system.

(6) Adaptability:
Processes in general change and evolve due to changes in external inputs or structural changes due to retrofitting and so on. In order to be useful and practical, the diagnostic system should be adaptable to changes.

(7) Explanation Facility:
A diagnostic system should also provide explanations on how a fault originated and propagated to the current situation. This is very important for building credibility with operators. One would like the system to not only justify why certain hypotheses were proposed but also explain why certain other hypotheses were not proposed.

(8) Modeling Requirements:
The amount of modeling required for system development is an important issue. For fast and easy deployment, the modeling effort should be as minimal as possible.

(9) Storage and Computational Requirements:
Usually, quick real-time solutions would require algorithms and implementations which are less computationally complex, but might entail high storage requirements. One would prefer a diagnostic system that is able to achieve a reasonable balance on these two competing requirements.

(10) Multiple Fault Identifiability:
The ability to identify multiple faults is an important but a difficult requirement. It is a difficult problem due to the interacting nature of most faults. In a general nonlinear system, the interactions would usually be synergistic and hence a diagnostic system may not be able to use the individual fault patterns to model the combined effect of the faults. On the other hand, enumerating and designing separately for various multiple fault combinations would become combinatorially prohibitive for large processes.

4. CLASSIFICATION OF DIAGNOSTIC APPROACHES

The two main components in a diagnostic classifier are: (i) the type of knowledge used and (ii) the type of diagnostic search strategy. Diagnostic search strategy is usually strongly dependent on the knowledge representation scheme which in turn is largely influenced by the kind of *a priori* knowledge available. Hence, the type of *a priori* knowledge used is the most important distinguishing feature in diagnostic systems. In this review, I classify diagnostic systems based on the *a priori* knowledge used.

The basic *a priori* knowledge that is needed for fault diagnosis is the set of failures and the relationship between the observations (symptoms) and the failures. A diagnostic system may have them explicitly (as in a table lookup), or it may be inferred from some source of domain knowledge. The *a priori* domain knowledge may be developed from a fundamental understanding of the process using first principles knowledge. Such knowledge is referred to as deep, causal or model-based knowledge (Milne, 1987). On the other hand, it may be gleaned from past experience with the process. This knowledge is referred to as shallow, compiled, evidential or process history-based knowledge.

The model-based *a priori* knowledge can be broadly classified as qualitative or quantitative. The model is usually developed based on some fundamental understanding of the physics of the process. In quantitative models, this understanding is expressed in terms of mathematical functional relationships between the inputs and outputs of the system. In contrast, in qualitative model equations, these relationships are expressed in terms of qualitative functions centered around different units in a process.

In contrast to the model based approaches where *a priori* knowledge about the model (either quantitative or qualitative) of the process is assumed, in process history based methods only the availability of large amount of suitably annotated historical process data is required. There are different ways in which this data can be transformed and presented as *a priori* knowledge to a diagnostic system. This is known as feature extraction. This can proceed as either quantitative or qualitative feature extraction. In quantitative feature extraction one can perform either a statistical or non-statistical feature extraction. This classification of diagnostic systems is shown in Figure 2.

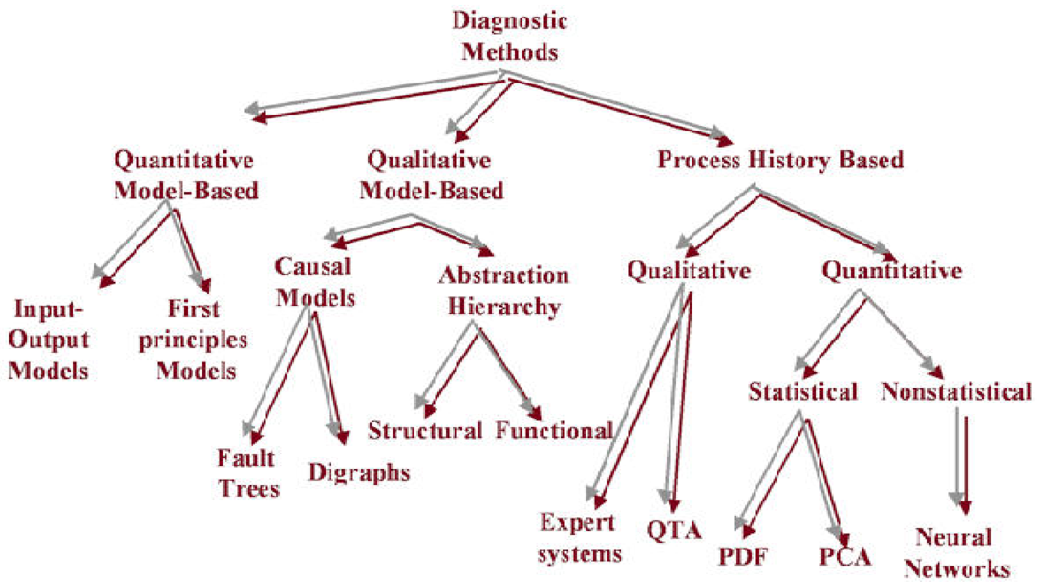

Fig. 2. Classification of diagnostic methods

5. QUANTITATIVE MODEL-BASED APPROACHES

This section briefly reviews quantitative model-based fault diagnosis methods. I briefly summarize and evaluate the most frequently-used fault detection and isolation approaches, such as observers, parity relations, Kalman filters and parameter estimation. For the sake of brevity, I have included only some of the many techniques available and therefore the references listed in this review paper are by no means exhaustive. Once again, for a more complete review the reader is referred to Venkatasubramanian et al. (2002a; 2002b; 2002c).

In automatic control, change/fault detection problems are known as model-based fault detection and isolation (FDI). Relying on an explicit model of the monitored plant, all model-based FDI methods (and many of the statistical diagnosis methods) require two steps. The first step generates inconsistencies between the actual and expected behavior. Such inconsistencies, also called *residuals*, are "artificial signals" reflecting the potential faults of the system. The second step chooses a decision rule for diagnosis.

The check for inconsistency needs some form of redundancy. There are two types of redundancies, hardware redundancy and analytical redundancy. The former requires redundant sensors. It has been utilized in the control of such safety-critical systems as aircraft, space vehicles and nuclear power plants. However, its applicability is limited due to the extra cost and additional space required. On the other hand, analytical redundancy is achieved from the functional dependence among the process variables and is usually provided by a set of algebraic or temporal relationships among the states, inputs and the outputs of the system (Chow and Willsky, 1984; Basseville, 1988; Frank, 1990). The essence of analytical redundancy is to check the actual system behavior against the system model for consistency. Any inconsistency, expressed as residuals, can be used for detection and isolation purposes. The residuals should be close to zero when no fault occurs but show "significant" values when the underlying system changes. To generate the diagnostic residuals requires an explicit mathematical model of the system. The model may be obtained either analytically using first principles or empirically as a black-box model. Also, statistical methods are often required for the decision making.

Although dynamic systems are continuous processes, all the diagnostic tools use sampled data, and hence only discrete models are included herein. However the basic concepts, if not the detailed analysis, carry over to continuous models. In addition, most of the model-based approaches assume linearity. Their application to a non-linear system requires a model linearization around the operating point.

Consider a system with m inputs and k outputs. Let $\mathbf{u}(t) = [u_1(t) \ldots u_m(t)]'$ be the process inputs and $\mathbf{y}(t) = [y_1(t) \ldots y_k(t)]'$ be the process outputs, where t denotes the discrete time. The basic system model in the state-space form is

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) \end{aligned} \tag{1}$$

where $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are parameter matrices with appropriate dimensions; $\mathbf{x}(t)$ denotes the n dimensional state vector.

The same system can be expressed in the input-output form

$$\mathbf{H}(z)\mathbf{y}(t) = \mathbf{G}(z)\mathbf{u}(t) \tag{2}$$

where $\mathbf{H}(z)$ and $\mathbf{G}(z)$ are polynomial matrices in z^{-1} (the backward-shift operator). $\underline{\mathbf{H}}(z)$ is diagonal; $\mathbf{H}(z)$ and $\mathbf{G}(z)$ are of the form

$$\begin{aligned} \mathbf{H}(z) &= \mathbf{I} + \mathbf{H_1}z^{-1} + \mathbf{H_2}z^{-2} + \ldots + \mathbf{H_n}z^{-n} \\ \mathbf{G}(z) &= \mathbf{G_0} + \mathbf{G_1}z^{-1} + \mathbf{G_2}z^{-2} + \ldots + \mathbf{G_n}z^{-n} \end{aligned}$$

Process models (1) and (2) describe an ideal situation where there are no faults or any form of disturbances and/or noise. Faults in the state-space framework are usually modeled by (Gertler, 1991; Gertler, 1992)

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{E}\mathbf{p}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t) + \mathbf{E}'\mathbf{p}(t) + \mathbf{q}(t) \end{aligned} \tag{3}$$

where input commands $\mathbf{u}(t)$ and measured outputs $\mathbf{y}(t)$ are both observable. Included in $\mathbf{p}(t)$ are actuator faults, certain plant faults, disturbances as well as input sensor faults. $\mathbf{q}(t)$ represents output sensor faults. In the input-output framework, (2) is replaced with

$$\mathbf{H}(z)\mathbf{y}(t) = \mathbf{G}(z)\mathbf{u}(t) + \mathbf{H}(z)\mathbf{q}(t) + \mathbf{F}(z)\mathbf{p}(t) \tag{4}$$

where $\mathbf{q}(t)$ and $\mathbf{p}(t)$ are as defined above.

For a detailed discussion on general diagnostic observer design for linear systems, the reader is referred to Frank (1990). Generation of diagnostic observers for nonlinear systems have also been considered to certain extent in the literature. An elegant approach for the generation of diagnostic observers for nonlinear systems which are in the fault-affine (similar to control-affine forms discussed in control literature) form can be found in (Frank, 1990). There have been other researchers who have looked at the problem of nonlinear observer design for a restricted class of nonlinear systems (Ding *et al.*, 1995; Yang and Saif, 1995).

Another approach to diagnosis is based on the parity equations notion. Parity equations are usually transformed variants of the input-output or state-space models of the plant (Gertler and Singer, 1990; Gertler, 1991). The essence is to check the parity (consistency) of the plant models with sensor outputs (measurements) and known process inputs. Under ideal steady-state operating conditions, the so-called residual or the value of the parity equations is zero. In real situations, the residuals are non-zero due to measurement and process noise, model inaccuracies, gross errors in sensors and actuators, and faults in the plant. The idea of this approach is to rearrange the model structure so as to get the best fault isolation. Dynamic parity relations were introduced by Willsky and Jones (1976). Further developments have been made by Gertler and Singer (1990), Gertler *et al.* (1995) and Gertler and Monajemy (1995) among others. It should be noted that all these methods are limited to faults that do not include gross process parameter drifts, and none of them address the issue of significant uncertainties in multiplicative parametric faults. It has been shown that once the residual properties have been selected, all parity equation and observer based designs are fundamentally equivalent (Gertler, 1992).

Plant disturbances are random fluctuations and often only their statistical parameters are known. One solution (Willsky, 1976; Basseville, 1988) to the fault diagnosis problem in such systems entails monitoring the innovation process or the prediction errors. The objective is to design a state estimator with minimum estimation error. It involves the use of optimal state estimate, *e.g.*, the Kalman filter, which is designed on the basis of the system model in its normal operating mode. The Kalman filter in state-space model is equivalent to an optimal predictor for a linear stochastic system in the input-output model. The statistical analysis of Kalman filter was pioneered by Willsky and Jones (1976) and further explored by Basseville and Benveniste (1986) and Basseville and Nikiforov (1993) and the references therein.

Isermann (1984) and Young (1981) surveyed different parameter estimation techniques such as least squares, instrumental variables and estimation via discrete-time models. Park et al.

(1982,1983,1987) have discussed the problem of parameter estimation and fault detection and diagnosis via parameter estimation. These methods require the availability of accurate dynamic models of the process and are computationally very intensive for large processes. The important issue here is complexity. If the process model is a complex nonlinear first principles model, then the parameter estimation problem turns out to be a nonlinear optimization problem. Real-time solution to complex nonlinear optimization problems is a serious bottleneck in the application of such approaches. Reduced order or input-output data models could be used, but this raises robustness issues.

Most of the work on model-based diagnostic systems has so far been concentrated in the aerospace, mechanical and electrical engineering literature. There has not been much work on the application of observers for fault diagnosis in chemical process systems. This may be due to the fact that the objectives of state/parameter estimation techniques when used in process control are different from the objectives of fault diagnosis as pointed out by Watanabe *et al.* (1994). Further, the unavailability and/or the complexity of high fidelity models and the essential nonlinear nature of these models for chemical processes render the design of diagnostic observers for such systems quite difficult. Gertler *et al.* (1995) presented a design procedure to generate isolable parity equation models chosen from a multitude of suitable models on the basis of sensitivities with respect to different failures and robustness relative to uncertainties in selected parameters. They illustrated the application of the technique on a distillation column. The application to EKF-based FDI system for fault diagnosis in the Model IV FCCU case study involving DAEs was reported by Huang *et al.* (2000). The application of three unknown input observers (UIO), one a *linear*, second an *extended linear* and the third a *non-linear* UIO on a CSTR case study is discussed in Dash *et al.* (2001a; 2001b). In their work, the performance of these three observers are evaluated through extensive simulation studies.

The type of models the analytical approaches can handle are limited to linear, and in some cases, to very specific nonlinear models. For a general nonlinear model, linear approximations can prove to be poor and hence the effectiveness of these methods might be greatly reduced. Another problem in this approach is the simplistic approximation of the disturbances that include modeling errors. In most cases, the disturbance matrix includes only additive uncertainty. However, in practice, severe modeling uncertainties due to parameter drifts come in the form of multiplicative uncertainties. This is a general limitation of all the model-based approaches that have been developed so far. In addition to difficulties related to modeling, they do not support an explanation facility owing to their procedural nature. Further, *a priori* estimation of classification errors can not also be provided using these methods. Another disadvantage with these methods is that if a fault is not specifically modeled (novelty identifiability), there is no guarantee that the residuals will be able to detect it. Adaptability of these approaches to varying process conditions has also not been considered. When a large-scale process is considered, the size of the bank of filters can be very large increasing the computational complexity, though, with the recent increase in computational power and the essential linear nature of these problems, this might not be a serious bottle-neck.

6. QUALITATIVE MODEL-BASED METHODS

In this part of the paper, I briefly review qualitative model representations and search strategies used in fault diagnostic systems. Qualitative models are usually developed based on some fundamental understanding of the physics and chemistry of the process. Various forms of qualitative models such as causal models and abstraction hierarchies have been developed. In terms of search strategies, one can broadly classify them as topographic and symptomatic search techniques. Topographic searches perform malfunction analysis using a template of normal operation, whereas, symptomatic searches look for symptoms to direct the search to the fault location.

In the causal models approach, people have developed systems based on signed digraphs (Iri *et al.*, 1979; Venkatasubramanian and Rich, 1988), qualitative simulation (Kuipers, 1985), qualitative process theory (Forbus, 1984), causal or precedence ordering techniques (Iwasaki and Simon, 1986) and so on. There are two main problems with qualitative models: ambiguities and spurious solutions. Ambiguities can be resolved completely only through the use of actual quantitative values. Frameworks for reasoning about relative orders of magnitudes have been proposed by (Raiman, 1986; Mavrovouniotis and Stephanopoulos, 1987). In these frameworks, influence magnitudes are related using relations such as A is negligible compared to B, A is close to B and A is the same order of magnitude as B. A set of inference rules then generates a partial ordering of values into groups significantly different in magnitude (Ungar and Venkatasubramanian, 1990). Spurious solutions refer to the generation of physically unrealizable solutions by a qualitative process. This problem can be alleviated to a reasonable extent by modeling the system from different perspectives (Kuipers, 1985; Kay and Kuipers, 1993).

Another form of qualitative modeling is the development of abstraction hierarchies based on decomposition of the system. The idea of decomposition is to be able to draw inferences about the behavior of the overall system solely from the laws governing the behavior of its subsystems. In such a decomposition, the no-function-in-structure principle is central: the laws of the subsystem may not presume the functioning of the whole system (de Kleer and Brown, 1984). In a hierarchic description, one could represent a generic description of a class of process units. The governing equations describing an entire class of process units may make assumptions about the class as a whole but may not make any assumptions about the behavior of particular units. Another important principle for decomposition of systems is the principle of locality: the laws for a part specifically cannot refer to any other part. No-function-in-structure allows consistent behaviors among various units. Principle of locality permits the behavior to be predicted based only on local information.

There are two-dimensions along which abstraction at different levels is possible - structural and functional (Rasmussen, 1986). The structural hierarchy represents the connectivity information of the system and its subsystems. The functional abstraction hierarchy represents the means-end relationships between a system and its subsystems. Majority of the work in fault diagnosis in chemical engineering depends on the development of functional decomposition. Structural decomposition is an efficient decomposition in systems where there is a general equivalence between structure and function, like for example in an electrical circuit. The reason for the popularity of functional decomposition in chemical engineering is due to the complex functionalities of various units that cannot be expressed in terms of structure. Hence, the decomposition focused here is the functional decomposition.

Diagnosis can be considered as a top-down search from a higher-level abstraction where groups of equipment and functional systems are considered to a lower-level of abstraction where individual units and unit functions are analyzed (Rasmussen, 1985). Based on this understanding, Shum and Davis (1985) decomposed the process into a hierarchy of functional subsystems. Each node in the hierarchy corresponds to the intended function of a subsystem. By comparing the function of the subsystem with the intended function, the hypothesis that a fault is present in that subsystem is evaluated. Finch and Kramer (1987) represent the plant as a set of interacting subsystems, where each subsystem is categorized as a control system (closed loops) or passive system (open loops) or an external system. Each of these subsystems has an associated function at this level of system description. Depending on their function, these subsystems are categorized as: (i) Functional, stressed, uncontrolled or saturated in the case of Control Systems and (ii) Functional or malfunctional in the case of Passive and External Systems.

Hierarchic abstraction of mass and energy flow structures at different levels of function (called Multilevel Flow Models), *i.e.*, functional abstraction hierarchy, has been used by Lind (1991). An MFM model is a normative description of a system, a representation of its function in terms of what should be done, how should it be done and with what should it be done. This leads to three basic concepts in MFM: (i) Goals, (ii) Functions and (iii) Physical components. Three types of connectional relations such as (i) achieve relations, (ii) condition relations and (iii) realize relations are used to connect objects. Diagnostic reasoning strategies based on the MFM model can be found in (Larsson, 1994).

Though qualitative models have a number of advantages as noted above, the major disadvantage is the generation of spurious solutions. Considerable amount of work has been done in the reduction of spurious solutions while reasoning with qualitative models. In SDGs, this is done using generation of latent constraints and similar techniques have been proposed for qualitative physics based models such as QSIM. The search strategies can be classified as either topographic or symptomatic search. Clearly, for a given qualitative representation, different search strategies could be used for diagnosis. Hence, one can view the methods proposed in the literature as different combinations of the qualitative models and search strategies.

7. PROCESS HISTORY-BASED METHODS

In this section, fault diagnosis methods based on historic process knowledge and data are reviewed. In contrast to the model based approaches where *a priori* knowledge (either quantitative or qualitative) about the process is needed, in process history based methods, a large amount of historical process data that is suitably annotated is needed. There are different ways in which this data can be transformed and presented as *a priori* knowledge to a diagnostic system. This is known as feature extraction. This extraction process can be either qualitative or quantitative in nature. Two of the major methods that extract qualitative history information are the expert systems (Davis, 1984; Rich and Venkatasubramanian, 1987) and trend modeling methods (Cheung and Stephanopoulos, 1990; Bakshi and Stephanopoulos, 1992; Rengaswamy and Venkatasubramanian, 1995; Vedam and Venkatasubramanian, 1997). Methods that extract quantitative information can be broadly classified as non-statistical or statistical methods. Neural networks are an important class of non-statistical classifiers (Leonard and Kramer, 1990; Kavuri and Venkatasubramanian, 1994). PCA/PLS and statistical pattern classifiers form a major component of statistical feature extraction methods (Kramer, 1991; Nomikos and MacGregor, 1994; Dong and McAvoy, 1996; Dunia *et al.*, 1996).

It is impossible to adequately review the abundant literature on expert systems for fault diagnosis here. A more detailed review of this is presented elsewhere (Venkatasubramanian *et al.*, 2001c). Most of the successful expert systems applications have been rule-based systems for structured selection or heuristic classification perspective of diagnosis, particularly for medical diagnosis. They often use some kind of probabilistic reasoning such as certainty factors, Bayesian methods or fuzzy logic to handle the uncertainties. Such systems, while quicker to develop and implement, are of limited scope, valid only for the processes for which they were developed.

In the trend representation approach, Cheung and Stephanopoulos (1990) presented a formal framework for the representation of process trends. They introduce triangulation to represent trends. Janusz and Venkatasubramanian (1991) propose a comprehensive set of primitives using which any trend can be represented. They use a finite difference method to calculate the first and second derivative of the process trend changes and based on these values, the primitives are identified. This qualitative formalism readily lends itself to hierarchic representations as well. Rengaswamy and Venkatasubramanian (1995) have shown how primitives can be extracted from raw noisy sensor data by treating the problem of primitive identification as a classification problem using neural networks. Vedam and Venkatasubramanian (1997) proposed a wavelet theory based adaptive trend analysis framework and later proposed a dyadic B-Spline based trend analysis algorithm (Vedam *et al.*, 1998). Recently, Rengaswamy *et al.* (2001) have discussed the utility of trend modeling in control loop performance assessment.

Multivariate statistical methods such as PCA and PLS have been used in diagnosis with considerable success (Qin and McAvoy, 1992; Nomikos and MacGregor, 1994; Dong and McAvoy, 1996; Vedam and Venkatasubramanian, 1997;Yoon and MacGregor, 2001). Overviews of using PCA and PLS in process analysis and control, fault detection and diagnosis were given by MacGregor *et al.* (1991; 1994) and MacGregor and Kourti (1995).

In an earlier work, Kresta *et al.* (1991) laid out the basic methodology of using multivariate statistical process control procedure to handle large numbers of process and quality variables for continuous process. Later on, Nomikos and MacGregor (1994) extended the use of multivariate projection methods to batch processes by using multiway PCA. To deal with nonlinearity, Qin and McAvoy (1992) proposed a neural net PLS approach that incorporated feedforward networks into the PLS modeling. In order to handle nonlinearity in batch processes, Dong and McAvoy (1996) utilized a nonlinear principal component analysis method. To facilitate the diagnosis procedure in very large processes, new hierarchical multivariate monitoring methods based on multiblock PLS algorithm was presented by MacGregor *et al.* (1991; 1994). Raich and Cinar (1996) proposed an integral statistical methodology combining principal component analysis and discrimination analysis techniques. Based on angle discriminants, a novel disturbance diagnosis approach (Raich and Cinar, 1997) was later introduced that provides better results for cases in which distance based discrimination is not accurate enough. Recently, Yoon and MacGregor (2000) have discussed the use of statistical and causal model-based methods for fault diagnosis.

Considerable interest has been shown in the literature in the application of neural networks for the problem of fault diagnosis (Himmelblau, 1986; Venkatasubramanian and Chan, 1989; Kavuri and Venkatasubramanian, 1994; Bulsari, 1995). Neural networks have been proposed for classification and function approximation problems. In general, neural networks that have been used for fault diagnosis can be classified along two directions: (i) the architecture of the network such as sigmoidal, radial basis and so on and, (ii) the learning strategy such as supervised and unsupervised learning. The most popular supervised learning strategy in neural networks has been the back-propagation algorithm. There are a number of papers that address the problem of fault detection and diagnosis using back-propagation neural networks. In chemical engineering, Venkatasubramanian (1985), Watanabe *et al.* (1989) and Venkatasubramanian and Chan (1989), Ungar *et al.* (1990) and Hoskins *et al.* (1991) were among the early researchers to demonstrate the usefulness of neural networks for fault diagnosis. Later, a more detailed and thorough analysis of the learning, recall and generalization characteristics of neural networks was presented by Venkatasubramanian *et al.* (1990) and Vaidhyanathan and Venkatasubramanian (1992). An hierarchical neural network architecture for the detection of multiple faults was proposed by Watanabe *et al.* (1994).

8. A COMPARISON OF VARIOUS APPROACHES

I have reviewed so far the three conceptually different frameworks for process fault diagnosis. In this section, I provide a comparative evaluation of these different frameworks against a common set of desirable characteristics for a diagnostic system that we proposed in section 3. The evaluations are summarized in Table 1. It is clear from the table that no single method is adequate to handle all the requirements for a desirable diagnostic system. Though all the methods are restricted, in the sense that they are only as good as the quality of information provided, it is seen that some methods might better suit the knowledge available than others. It is our view that some of these methods can complement one another resulting in better diagnostic systems. Integrating these complementary features is one way to develop hybrid methods that could overcome the limitations of individual solution strategies. Hence, hybrid approaches where different methods work in conjunction to solve parts of the problem are attractive.

Combination of methods allows one to evaluate different kinds of knowledge in one single framework for better decision making. A blackboard-based cooperative problem-solving framework where different diagnostic methods work in conjunction to perform collective process fault diagnosis has been proposed by (Mylaraswamy *et al.*, 1994; Mylaraswamy and Venkatasubramanian, 1997; Vedam and Venkatasubramanian, 1999). The blackboard architecture, called Dkit, in which different diagnostic methods analyze the same problem, and a scheduler which regulates the decision-making of these methods, is the central concept in this framework. The utility of such a hybrid framework for solving real-time complex fault diagnosis problems is illustrated through the use of a diagnosis study on the Amoco Model IV Fluid Catalytic Cracking Unit (FCCU). This framework was adopted by the Honeywell ASM Consortium for the development of a commercially-viable, next generation, intelligent control systems called AEGIS and MSEP.

Table 1. Comparison of various diagnostic methods

	Observer	Digraphs	Abstraction Hierarchy	Expert Systems	QTA	PCA	Neural Networks
Quick Detection and Diagnosis	✓	?	?	✓	✓	✓	✓
Isolability	✓	×	×	✓	✓	×	✓
Robustness	✓	✓	✓	✓	✓	✓	✓
Novelty Identifiability	?	✓	✓	×	?	✓	✓
Classification Error	×	×	×	×	×	×	×
Adaptability	×	✓	✓	×	?	×	×
Explanation Facility	×	✓	✓	✓	✓	×	×
Modelling Requirement	?	✓	✓	✓	✓	✓	✓
Storage & Computation	✓	?	?	✓	✓	✓	✓
Multiple Fault Identifiability	✓	✓	✓	×	×	×	×

✓ FAVORABLE
X NOT FAVORABLE
? SITUATION DEPENDENT
— NOT APPLICABLE

9. Intelligent Systems for Process Hazards Analysis

As noted earlier, AEM and PHA are closely related activities that deal with process safety and operability issues and hence have an impact on supervisoy control decisions. Again, concerns about process safety have led the federal agencies in the U.S. to create stricter safety, health and environmental regulations that require the performance of process hazards analysis. A typical HAZOP analysis can take 1-8 weeks to complete, costing over $13,000-25,000 per week. By an OSHA estimate, approximately 25,000 plant sites in the United States require a PHA (Freeman et al., 1992). An estimated $5 billion is spent annually by the chemical process industries (CPI) to perform PHAs and related activities. The estimated cost of process hazards reviews in the CPI is about 1% of sales or about 10% of profits.

Given the enormous amounts of time, effort and money involved in performing PHA reviews, there exist considerable incentives to develop intelligent systems for automating the process hazards analysis of chemical process plants. An intelligent system can reduce the time, effort and expense involved in a PHA review, make the review more thorough, detailed, and consistent, minimize human errors, and free the team to concentrate on the more complex aspects of the analysis which are unique and difficult to automate. Also, an intelligent PHA system can be integrated with CAD systems and used during early stages of design, to identify and decrease the potential for hazardous configurations in later design phases where making changes could be economically prohibitive. It would facilitate automatic documentation of the results of the analysis for regulatory compliance. Also these PHA results can be made available online to assist plant operators during diagnosis of abnormal situations as well as to train novice operators.

Despite the obvious importance of this area, there has only been limited work on developing intelligent systems for automating PHA of process plants. In this paper, we will review the past approaches towards the automation of PHA from the perspective of intelligent systems. This paper is written as a brief survey of the literature in this area with an emphasis on the overview of the results of the Purdue investigations on intelligent systems for PHA over the past 12 years. Of the various methods, HAZOP analysis is the most widely used and recognized as a preferred PHA approach by the chemical process industries. Hence, the main focus of this paper will be on HAZOP analysis. We will not provide a detailed review of HAZOP or the intelligent systems framework for it here as it is not the intent of this paper. It has been addressed elsewhere by the author (Venkatasubramanian et. al., 2000).

HAZOPExpert: A Model-Based Intelligent System for Continuous Processes

HAZOPExpert is a model-based, object-oriented, intelligent system for automating HAZOP analysis developed by Venkatasubramanian and Vaidhyanathan during 1990-94 for continuous processes. In their approach, they recognized that while the results of a HAZOP study may vary from plant to plant, the approach itself is systematic and logical, with many aspects of the analysis being the same and 'routine' for different process flowsheets. It turns out that about 70% of time and effort is spent on analyzing these 'routine' process deviations, their causes, and consequences. Hence, they focused on these 'routine' cause-and-effect analyses by developing generic models which can be used in a wide variety of flowsheets, thus making the expert system process-independent. They also recognized that the process-specific components of knowledge, such as the process material properties and process P&IDs, have to be flexibly integrated with the generic models in an appropriate manner. To address this integration, they developed a two-tier knowledge-based framework by decomposing the knowledge base into 'process specific' and 'process general' knowledge, represented in an object-oriented architecture.

Process-specific knowledge consists of information about the materials used in the process, their properties (such as corrosiveness, flammability, volatility, toxicity, etc.) and the P&ID of the plant. The process-specific knowledge is likely to change from plant to plant and is provided by the user. Process-general knowledge comprises of the process unit HAZOP models that are developed in a context-independent manner, which remain the same irrespective of the process plant under consideration. The HAZOP model of a process unit consists of its class definition and generic qualitative causal

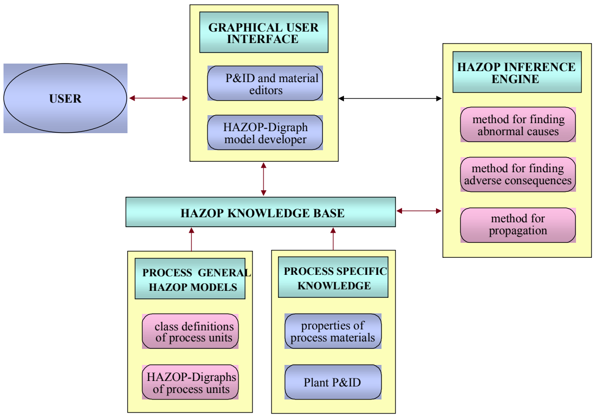

Fig. 3 Structure of HAZOPExpert

model-based methods for identifying and propagating abnormal causes and adverse consequences of process variable deviations. Based on this framework, an expert system called *HAZOPExpert* (Figure 3) has been implemented using Gensym's G2 real-time expert system shell. *HAZOPExpert's* inference engine allows for the interaction of the process-general knowledge with the process-specific knowledge to identify the valid abnormal causes and adverse consequences for the given process variable deviations for the particular HAZOP study section of the plant under consideration.

HAZOPExpert is not meant to replace the HAZOP team. It's objective is to automate the 'routine' aspects of the analysis as much as possible, thereby freeing the team to focus on more complex aspects of the analysis that can not be automated.

Batch *HAZOPExpert*: A Model-Based Intelligent System for Batch Process PHA

Batch HAZOPExpert (BHE) is a model-based intelligent system for automating HAZOP analysis for batch processes based on HAZOPExpert. It was first developed by Srinivasan and Venkatasubramanian (1998a, 1998b), and improved later by Zhao, Viswanathan and Venkatasubramanian (1998) by considering batch control, dynamic propagation of materials concentration and quantitative process information.

For a batch process, HAZOP analysis is more complex because of the discrete-event character which raises issues about its temporal nature. The status of the plant is constantly changing in some established discrete sequence. Since the P&ID does not sufficiently define the system, a set of operating instructions and some form of sequence chart are also needed. In batch plants, these sequence and instructions are called product recipe. Product recipe consists of a series of tasks occurring at discrete instants of time. In each task, many subtasks are executed to achieve the task. Differential equation based mathematical model, therefore, is not enough to describe batch chemical processes. Tools for describing discrete systems such as Petri nets (Peterson, 1981) are often used to represent batch processes (Srinivasan et. al., 1998a).

In BHE, the product recipe is represented with two-layered Petri Nets - Recipe Petri Net (RPN) and Task Petri Net (TPN). RPN indicates the sequence of the tasks while TPN demonstrates the sequence of the subtasks in each task. Associated with each subtask, there is a digraph model which qualitatively captures the general cause-effect relationships between the variables of the subtask. Currently, about 40 subtask digraph models have been established in the model library to cover most of batch operations such as heat, cool, and filtration and so on.

Batch HAZOPExpert is the only intelligent system that has been successfully applied to batch chemical plants. It has been tested on more than twenty industrial batch processes with our industrial partners. More results on industrial applications of BHE can be found in the literature (Zhao et al., 2000).

Recently, these two systems have been combined into a single entity called the PHASuite. Using PHASuite has been found to result in savings of about 50% in time and money for typical industrial applications.

10. CONCLUSIONS AND FUTURE DIRECTIONS

Process safety, occupational health and environmental issues are among the top priorities of all process industries. The CPI collectively spend billions of dollars every year on AEM and PHA. These concerns can be effectively addressed by the emerging intelligent control systems paradigm which has the goal of automating supervisory control tasks currently handled manually by human operators. The basic aim of this paper was to give a broad overview of the various approaches to automated fault diagnosis and describe the state-of-the-art efforts in terms of industrial applications in the field. Intelligent control components such as diagnostic methods based on model-based and process-history-based approaches were discussed. It was shown that no single method was capable of meeting all the requirements of a good diagnostic system and a hybrid framework involving collective problem solving and parallel ways of reasoning was recognized as an attractive alternative to address the challenges of complex, industrial-scale diagnostic problems. Qualitative and quantitative methods offer their inherent advantages respectively and a combination of these would certainly be suited to address the challenges at hand (Frank, 1990).

This paper also briefly reviewed how manual process hazards analysis techniques and methodologies can be automated by an intelligent systems approach to reduce the time, effort and cost involved, and to improve the consistency and thoroughness of the analysis. Such systems are not meant to replace the human team but to assist them in improving the overall efficiency and productivity of the team. The HAZOPexpert, Batch HAZOPexpert and PHASuite intelligent systems developed at Purdue University are now well beyond proof of concept and are ready for industrial applications and commercial exploitation.

In addition to routine PHA, such intelligent systems can facilitate HAZOP reviews at an early stage of process development and design. This means that problems can be identified and rectified during detailed design or while formulating operating procedures. Making changes once a plant is built are very expensive compared with changes at the design stage (Skelton, 1997). Early identification of hazards will also lead to effective avoidance or control of such hazards. HAZOP at this stage will also help to develop confidence that the desired process is safe. Along these lines, the longer-term aim may well be to move towards process conception and synthesis to create inherently safer designs and operating plants that tend towards zero defects. A more immediate development could be the use of online hazard reviews for the training of operators for abnormal situation management. The online hazard models can also be adapted for fault diagnosis applications.

As one can gather from the observations in this paper, integrating the complementary methods in a synergistic fashion to perform collective problem solving is key to future improvements in diagnostic systems. Also one has to look at the relevance of diagnosis to other process operations (Venkatasubramanian and Stanley, 1993). Integration of diagnosis with other process operations like data reconciliation and regulatory control can help solve the problem of process operations management effectively. Progress in real-time dynamic simulation and development of more accurate models of chemical processes would allow the application of model-based techniques, bringing in their inherent advantages to the goal of situation assessment and rectification. Efforts also are being made to utilize the results of Process Hazards Analysis (PHA) – carried out *off-line* - in AEM (Dash and Venkatasubramanian, 1999).

The application of the intelligent systems framework for complex problems such as abnormal situation management and process hazards analysis, which were formerly solved only by human teams, has come a long way since its modest beginnings in the 1980s. Intelligent systems are now well poised to make significant contributions to AEM and PHA in real-life industrial settings thus revolutionizing process control in the coming decade in a wide variety of industries.

REFERENCES

ASM Consortium Web Site: http://www.iac.honeywell.com/pub/absitmang/

Bailey, S. J. (1984). From desktop to plant floor, a CRT is the control operators window on the process. *Control Engineering*, **31**(6), 86-90.

Bakshi, B. and G. Stephanopoulos (1992). Temporal representation of process trends for diagnosis and control. In: *IFAC Symposium on Online Fault Detection and Supervision in the Chemical Process Industries*.

Basseville, M. and A. Benveniste (Eds.) (1986). *Detection of abrupt changes in signals and dynamic systems (Lecture Notes in Control and Information Sciences: 77)*. Springer-Verlag, Berlin.

Basseville, M. (1988). Detecting changes in signals and systems. *Automatica*, **24** (3), 309-326.

Basseville, M. and I. V. Nikiforov (1993). *Detection of Abrupt Changes-Theory and Application*. Prentice Hall, Information and System Sciences Series.

Bulsari (Editor), A. B. (1995). *Neural networks for chemical engineers*. Elsevier Science, Amsterdam.

Bureau of Labor Statistics (1998). Occupational injuries and illnesses in the United States by industry. Government Printing Office, Washington, DC.

Cheung, J. T. and G. Stephanopoulos (1990). Representation of process trends –I: A formal representation framework. *Comput. & Chem. Engng.,* **14**, 495-510.

Chow, E. D. and A. S. Willsky (1984). Analytical redundancy and the design of robust failure detection systems. *IEEE Trans. Automat. Contr.* **29** (7), 603-614.

Dash, S. and V. Venkatasubramanian (1999). An Integrated framework for Abnormal situation management and Process hazards analysis. In: *AIChE Annual Meeting,* Dallas.

Dash, S., S. Kantharao, R. Rengaswamy and V. Venkatasubramanian (2001). Application and Evaluation of a Linear/Restricted Nonlinear Observer to a Nonlinear CSTR. *European Symposium on Computer Aided Process Engineering – 11*, Kolding, Denmark, 853-858.

Dash, S., P. Vachhani, R. Rengaswamy and V. Venkatasubramanian (2002). Applications of observers for fault diagnosis in a nonlinear CSTR. To be submitted to *Chemical Engineering Science*.

Davis, R. (1984). Diagnosis reasoning based on structure and behavior. *Artif. Intell.* **24**, 347-410.

de Kleer, J. and S. Brown (1984). A qualitative physics based on confluences. *Artificial Intelligence*. **24**, 7-83.

Ding, Y., J. B. Gomm, D. N. Shields, D. Williams and K. Disdell (1995). Fault diagnosis for a gas-fired furnace using bilinear observer method. In: *Proceedings of the American Control Conference, Seattle, Washington*.

Dong, D. and T. J. McAvoy (1996). Batch tracking via nonlinear principal component analysis. *AIChE J.,* **42** (8), 2199-2208.

Dunia, R., S. J. Qin, T. F. Edgar and T. J. McAvoy (1996). Identification of faulty sensors using principal component analysis. *AIChE J.*, **42**(10), 2797–2812.

Finch, F. E. and M. A. Kramer (1987). Narrowing diagnostic focus using functional decomposition. *AIChE J.*, **34** (1), 130-140.

Forbus, K. D. (1984). Qualitative process theory. *Artif. Intell.*, **24,** 85-168.

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

Freeman, R. A., R. Lee and T. P. McNamara (1992). Plan HAZOP studies with an expert system. *Chem. Engng. Prog.*, **88**(8), 28-32.

Gertler, J. and D. Singer (1990). A new structural framework for parity equation-based failure detection and isolation. *Automatica*, **26**, 381-388.

Gertler, J. (1991). Analytical redundancy methods in fault detection and isolation. In: *Proc. IFAC/IMACS Symp. SAFEPROCESS'91, Baden-Baden.*

Gertler, J. (1992). Analytical redundancy methods in fault detection and isolation - survey and synthesis. In: *IFAC symposium on Online Fault Detection and Supervision in the Chemical Process Industries*.

Gertler, J., M. Costin, X. Fang, Z. Kowalczuk, M. Kunwer and R. Monajemy (1995). Model based diagnosis for automative engines- algorithm development and testing on a production vehicle.

IEEE Transactions on Control Systems Technology, **3**, 61-69.

Gertler, J. and R. Monajemy (1995). Generating directional residuals with dynamic parity relations. *Automatica*, **31**, 627-635.

Himmelblau, D. M. (1978). Fault detection and diagnosis in chemical and petrochemical processes. , Elsevier Press, Amsterdam.

Himmelblau, D. M. (1986). Fault detection and diagnosis - today and tomorrow. In: *IFAC Kyoto Workshop on fault detection and Safety in Chemical Plants*. Kyoto, Japan.

Hoskins, J. C., K. M. Kaliyur and D. M. Himmelblau (1991). Fault diagnosis in complex chemical plants using artificial neural networks. *AIChE J.,* **37**, 137-141.

Huang, Y., S. Dash, G. V. Reklaitis and V. Venkatasubramanian (2000). EKF based estimator for FDI in model IV FCCU. In: *SAFEPROCESS 2000, 14-16 June*. Budapest, Hungary.

Iri, M., K. Aoki, E. O'Shima and H. Matsuyama (1979). An algorithm for diagnosis of system failures in chemical processes. *Comput. & Chem. Engng.*, **3**, 489-493.

Isermann, R. (1984). Process faults detection based on modeling and estimation methods - a survey. *Automatica*, **20** (4), 387-404.

Iwasaki, Y. and H. Simon (1986). Causality in Device Behavior., *Artificial Intelligence*, **29**, 3-32.

Janusz, M. and V. Venkatasubramanian (1991). Automatic generation of qualitative description of process trends for fault detection and diagnosis. *Engng. Applic. Artif. Intell.*, **4**, 329-339.

Kavuri, S. N. and V. Venkatasubramanian (1994). Neural network decomposition strategies for large-scale fault diagnosis. *Int. J. Control*, **59**(3), 767-792.

Kay, H. and B. Kuipers (1993). Numerical behavior envelopes for qualitative models. In: *Proceedings of AAAI-93*, pp. 606-613.

Kletz, T. A. (1988). What went wrong? - case histories of process plant disasters. Chapter **18**, 189-196, 2nd edition, Gulf Publishing, Houston, Texas.

Kletz, T. A. (1991). Incidents that could have been prevented by HAZOP. *Journal of Loss Prev. Process Ind*., **4**, 128-129.

Kramer, M. A. (1991). Principal component analysis using auto-associative neural networks. *AIChE J.*, **37** (2), 233-243.

Kresta, J. V., J. F. MacGregor and T. E. Marlin (1991). Multivariate Statistical Monitoring of Processes. *Can. J. Chem. Eng.*, **69,** 35.

Kuipers, B. (1985). The limits of qualitative simulation. In: *Proceedings of Ninth Joint International Conference on Artificial Intelligence*.

Larsson, J. E. (1994). Diagnostic reasoning strategies for means-end models. *Automatica*, **30** (5), 775-787.

Laser, M. (2000). Recent safety and environmental legislation. *Trans IchemE* **78**(B), September, 419-422.

Lees F. P. (1996). Loss prevention in the process industries: hazard identification, assessment and control. Second edition, Butterworth-Heinemann, British.

Leonard, J. A. and M. A. Kramer (1990). Limitations of backpropagation approach to fault diagnosis and improvements with radial basis functions. In: *AIChE Annual Meeting, Chicago*.

Lind, M. (1991). Abstraction for modeling diagnostic strategies. In: *IFAC Workshop on Computer Software Structures Integrating AIKBS Systems in Process Control*.

MacGregor, J. F., T. E. Marlin, J. Kresta and B. Skagerberg (1991). Multivariate statistical methods in process analysis and control. In: *Chemical Process Control - CPCIV* (Y. Arkun and W. H. Ray, Eds.). pp.79-100. CACHE-AIChE.

MacGregor, J. F., J. Christiana K. Costas and M. Kotoudi (1994). Process Monitoring and diagnosis by multiblock PLS methods. *AIChE J.*, **40**(5), 826-838.

MacGregor, J. F. and T. Kourti (1995). Statistical process control of multivariate processes. *Control Engng. Practice*, **3** (3), 403-414.

Mavrovouniotis, M. and G. Stephanopoulos (1987). Reasoning with order of magnitudes and approximate relations. In: *Proceedings of AAAI-87, July*.

Milne, R. (1987). Strategies for diagnosis. *IEEE Trans. Syst., Man, and Cybernetics,* **17**(3), 333-339.

Mylaraswamy, D. (1996). Dkit: A blackboard-based distributed multi-expert environment for abnormal situation management. *PhD Thesis*, Purdue University.

Mylaraswamy, D. and V. Venkatasubramanian (1997). A hybrid framework for large scale process fault diagnosis. *Comput. & Chem. Engng* **21**, S935-S940.

National Safety Council (1999). Injury Facts 1999 Edition. National Safety Council, Chicago.

Nimmo, I. (1995). Adequately address abnormal situation operations. *Chem. Engng. Prog.*, **91**(9), 36-45.

Nomikos, P. and J. F. MacGregor (1994). Monitoring of batch processes using multiway principal component analysis. *AIChE J.*, **40**, 1361-1375.

OSHA (1992). Process safety management of highly hazardous chemicals; explosives and blasting agents; final rule. 29 CFR 1910.119: Federal Register, February 24, 6356-6417.

Park, S. W., and D. M. Himmelblau (1982). Parameter Estimation and Identifiability. Chem Eng J Bioch Eng, **25**(2), 163-174.

Park, S. W., and D. M. Himmelblau (1983). Fault Detection and Diagnosis via Parameter Estimation in Lumped Parameter Systems. Industrial and Engineering Chemistry, Proc. Des. and Dev. **22**, 482-487.

Park, S. W., and D. M. Himmelblau (1987). Structural Design for Systems Fault Diagnosis. *Comput. & Chem. Engng* , 11(6), 713-722.

Peterson, J. L. (1981). Petri net theory and the modeling of systems. New Jeresy: Prentice-Hall.

Qin, S. J. and T. J. McAvoy (1992). Nonlinear pls modeling using neural networks. *Comput. & Chem. Engng*., **16** (4), 379-391.

Raich, A. and A. Cinar (1996). Statistical process monitoring and disturbance diagnosis in multivariable continuous processes. *AIChE J.*, **42**, 995-1009.

Raich, A. and A. Cinar (1997). Diagnosis of process disturbances by statistical distance and angle measures. *Comput. & Chem. Engng.*, **21** (6), 661-673.

Raiman, O. (1986). Order of magnitude reasoning. In: *Proceedings of AAAI-86, August*.

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decision making and system management. *IEEE Trans. Syst., Man, and Cyber.*, **15** (2), 234-245.

Rasmussen, J. (1986). *Information Processing and Human-machine Interaction*. North Holland, New York.

Rengaswamy, R. and V. Venkatasubramanian (1995). A syntactic pattern recognition approach for process monitoring and fault diagnosis. *Engng. Applic. Artif. Intell.*, **8**(1), 35-51.

Rengaswamy, R., T. Hagglund and V. Venkatasubramanian (2001). A qualitative shape analysis formalism for monitoring control loop performance. *Engng. Applic. Artif. Intell.*, **14**(1), 23-33.

Rich, S. H. and V. Venkatasubramanian (1987). Model-based reasoning in diagnostic expert systems for chemical process plants. *Comput. & Chem. Engng.*, **11** (2), 111-122.

Shum, S. K. and J. F. Davis (1985). An expert system for diagnosing process plant malfunctions. In: *IFAC Workshop on Fault Detection and Safety in Chemical Plants, Kyoto, Japan*.

Skelton, B. (1997). Process safety analysis: an introduction. Houston: Gulf Publishing.

Srinivasan, R. and V. Venkatasubramanian (1998a). Automating HAZOP analysis of batch chemical plants: Part I. The knowledge representation framework. *Comput. & Chem. Engng.*, **22**(9), 1345-1355.

Srinivasan, R. and V. Venkatasubramanian (1998b). Automating HAZOP analysis of batch chemical plants: Part II. Algorithms and applications. *Comput. & Chem. Engng.*, **22**(9), 1357-1370.

Ungar, L. H., B. A. Powell and S. N. Kamens (1990). Adaptive networks for fault diagnosis and process control. *Comput. & Chem. Engng.* **14**, 561-573.

Ungar, L. H. and V. Venkatasubramanian (1990). *Artificial Intelligence in Process Systems Engineering: Knowledge Representation*. CACHE, Austin, Texas.

Vaidyanathan, R. and V. Venkatasubramanian (1992). Representing and diagnosing dynamic process data using neural networks. *Engng. Applic. Artif. Intell.*, **5** (1), 11-21.

Vedam, H. and V. Venkatasubramanian (1997). A wavelet theory-based adaptive trend analysis system for process monitoring and diagnosis. In *American Control Conference*, 309-313.

Vedam, H., V. Venkatasubramanian and R. Bhalodia (1998). A B-Spline based method for data compression, process monitoring and diagnosis. *Comput. & Chem. Engng.*, **22**, S827-S830.

Vedam, H., S. Dash and V. Venkatasubramanian (1999). An intelligent operator decision support system for abnormal situation management. *Comput. & Chem. Engng.*, 23S, S577-S580.

Vedam, H., and V. Venkatasubramanian (1999). PCA-SDG based process monitoring and fault diagnosis. *Control Engng. Practice* **7** (7), 903-17.

Venkatasubramanian, V. (1985). Inexact reasoning in expert systems: a stochastic parallel network approach. In: *Second Conference on Artificial Intelligence Applications: The Engineering of Knowledge-Based Systems*. IEEE Comput. Soc. Press., Washington, DC, USA. pp. 13-15.

Venkatasubramanian, V. and S. H. Rich (1988). An object-oriented two-tier architecture for integrating compiled and deep-level knowledge for process fault diagnosis. *Comput. & Chem. Engng.*, **12** (9), 903-921.

Venkatasubramanian, V. and K. Chan (1989). A neural network methodology for process fault diagnosis. *AIChE J*.**,** **35**, 1993-2002.

Venkatasubramanian, V., R. Vaidyanathan and Y. Yamamoto (1990). Process fault detection and diagnosis using neural networks I: Steady state processes. *Comput. & Chem. Engng.*, **14** (7), 699-712.

Venkatasubramanian, V. and G. M. Stanley (1993). Integration of process monitoring, diagnosis and control: issues and emerging trends. In: *Proceedings of the second International Conference on Foundations of Computer Aided Operations,* Colorado.

Venkatasubramanian, V., J. Zhao, and S. Viswanathan (2000). Intelligent systems for HAZOP analysis of complex process plants. *Comput. & Chem. Engng.*, **24**(9-10), 2291-2302.

Venkatasubramanian, V., R. Rengaswamy, K. Yin and S. N. Kavuri (2002a). A review of process fault detection and diagnosis Part I: quantitative model based methods. To appear in *Comput. & Chem. Engng.*

Venkatasubramanian, V., R. Rengaswamy and S. N. Kavuri (2002b). A review of process fault detection and diagnosis Part II: qualitative models and search strategies. To appear in *Comput. & Chem. Engng.*

Venkatasubramanian, V., R. Rengaswamy, S. N. Kavuri and K. Yin (2002c). A review of process fault detection and diagnosis Part III: process history based methods. To appear in *Comput. & Chem. Engng.*

Watanabe, K., I. Matsura, M. Abe, M. Kubota and D. M. Himmelblau (1989). Incipient fault diagnosis of chemical processes via artificial neural networks. *AIChE J.*, **35** (11), 1803-1812.

Watanabe, K., S. Hirota, L. Hou and D. M. Himmelblau (1994). Diagnosis of multiple simultaneous fault via hierarchical artificial neural networks. *AIChE J.*, **40** (5), 839-848.

Willsky, A. S. (1976). A survey of design methods for failure detection in dynamic systems. *Automatica*, **12**, 601-611.

Willsky, A. S. and H. L. Jones (1976). A generalized likelihood ratio approach to detection and estimation of jumps in linear systems. *IEEE Trans. on Automatic Control*, **AC-21**, 108-112.

Yang, H., and M. Saif (1995). Nonlinear adaptive observer design for fault detection. In: *Proceedings of the American Control Conference, Seattle, Washington*.

Yoon, S. and J. F. MacGregor (2000). Statistical and Causal Model-based Approaches to Fault Detection and Isolation. *AIChE J.,* **46**(9), 1813-1824.

Yoon, S. and J. F. MacGregor (2001). Fault Diagnosis with Multivariate Statistical Models. *Journal of Process Control*, **11**, 387-400.

Young, P. (1981). Parameter estimation for continuous time models - a survey. *Automatica*, **17** (1), 23-39.

Zhao. J., S. Viswanathan, V. Venkatasubramanian, J. Vinson, and P. Basu (1998). Automated process hazard analysis of batch chemical plants. In: *AIChE Annual Meeting*, Miami, USA.

Zhao, J., S. Viswanathan, and V. Venkatasubramanian (2000). Industrial applications of intelligent systems for operating procedure synthesis and hazards analysis for batch process plants. *Computer-Aided Chemical Engineering 8 (ESCAPE-10 Proceedings)*, 787-792.