Diagnosis of Continuous Valued Systems in Transient Operating Regions

Pieter J. Mosterman and Gautam Biswas, *Senior Member, IEEE*

Abstract— **The complexity of present day embedded systems (continuous processes controlled by digital processors), and the increased demands on their reliability motivate the need for monitoring and fault isolation capabilities in the embedded processors. This paper develops monitoring, prediction, and fault isolation methods for abrupt faults in complex dynamic systems. The transient behavior in response to those faults is analyzed in a qualitative framework using parsimonious topological system models. Predicted transient effects of hypothesized faults are captured in the form of *signatures* that specify future faulty behavior as higher order time-derivatives. The dynamic effects of faults are analyzed by a *progressive monitoring* scheme till transient analysis mechanisms have to be suspended in favor of steady state analysis. This methodology has been successfully applied to monitoring of the secondary sodium cooling loop of a fast breeder reactor.**

I. Introduction

THE complexity and sophistication of the new generation of engineered systems along with growing demands for their reliability, safety, and low cost operation, is being met by the use of more automated monitoring and fault detection and isolation (FDI) subsystems. The goal is to accurately[1] isolate problems and restore the system to normal operation by making control changes to bring system behavior back to desired operating ranges or at least a *safe* mode of operation. This defines a paradigm for fault detection, isolation, and recovery (FDIR).

Functional redundancy schemes use measured system variable values with relations imposed by the system configuration and functionality to analyze discrepancies among the measured values [20]. Deviations in measurement values can be expressed in terms of changed component parameter values, which are then mapped to faulty components. Traditional functional redundancy schemes employ state and parameter estimation methods, adaptive filtering, and logic based schemes for analysis [4]–[6].

System models capture relations between measured variables and system component parameters. FDI methods often employ failure models to establish the relations between measurements and a pre-enumerated set of faults [9], [18], but the disadvantage of this approach is that it fails to identify unusual and novel faults. More general functional models describe system behavior, and fault isolation is based on analysis of reported deviations in the context of the given model [1], [15]. Faults can be characterized as follows [4].

- Incipient faults occur slowly over time, and are linked to the wear and tear of components and drift in control parameters.
- Intermittent faults are only present for very short periods in time, but sometimes can have disastrous consequences.
- Abrupt faults are dramatic and persistent, and they cause significant deviations from steady state operations called *transients*. In time the system either moves into a new steady state or returns to its original steady state.

The difference in fault characteristics requires different schemes for effective and reliable detection and isolation. For example, parameter estimation methods which compute parameter values from input-output relations work well for incipient faults because the system changes slowly and tends to remain in steady state [5].

Our primary focus is on abrupt fault analysis in continuous dynamic systems. This makes it essential to track and analyze system behavior at frequent intervals from the point of failure so transient characteristics are not lost. Capturing behavior at or very close to the point of failure is important, because, as time progresses, compensating effects such as dynamic feedback may begin to mask the effects of the fault. Moreover, it may be impractical to rely on subsequent steady state analysis because the system may take a long time to reach a new steady state.

Fig. 1 illustrates a generic model based approach to fault detection and isolation [5], [6]. A set of variables, called *observations*, are monitored at frequent intervals. Deviations in observations imply faults.

Definition 1 (Observation): An observation is a variable in the system model that is measured.

Models that reason about dynamic behavior of the system are utilized to predict operating values for the chosen observations. Residuals, $\mathbf{r}$, are computed as the difference between the observations, $\mathbf{y}$, and predicted normal behavior, $\hat{\mathbf{y}}$ (Fig. 1). Nonzero residuals trigger the diagnosis algorithm. The analysis of these residuals in the context of the model *generates* one or more hypothesized *causes*, $\mathbf{f}$, that explain the observed deviations. System models with imposed hypothesized faults are then used to predict future system behavior.

Manuscript received August 10, 1997; revised November 4, 1998 and August 2, 1999. This work was supported in part under grants from PNC, Japan, and HP Labs, USA. This paper was recommended by Associate Editor J. M. Tien.

P. J. Mosterman is with the Institute of Robotics and System Dynamics, DLR Oberpfaffenhofen, D-82230 Wessling, Germany (e-mail: pieter.j.mosterman@dlr.de).

G. Biswas is with the Department of Computer Science, Vanderbilt University, Nashville, TN 37235 USA (e-mail: biswas@vuse.vanderbilt.edu).

Publisher Item Identifier S 1083-4427(99)08402-7.

[1] Accuracy is the assessment of whether the actual fault is contained by the estimate. Precision is a measure of the difference between the estimated faults and the true faults.

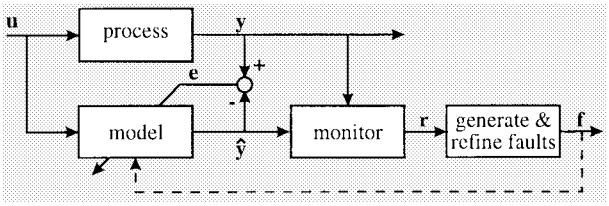


Fig. 1. Diagnosis of dynamic systems.

Continued monitoring and comparison with these predictions helps *refine* the initial fault set, $\mathbf{f}$. Faults whose predictions remain consistent with the observations establish the root-causes for the observed failures. Monitoring, comparison, and refinement continues till a unique fault has been isolated or transient analysis has to be suspended. The overall process of monitoring, hypothesizing faults, prediction, and fault isolation with explicit system models as the core of the analysis scheme is referred to as *model based diagnosis*.

In processes that operate mostly in steady state, nominal values and their upper and lower limits can often be retrieved from design specifications or documentation created by process engineers. For systems whose normal operation modes include transients and dynamic behaviors, it is harder to determine nominal values and thresholds from which deviations in process variables can be derived [19]. A fairly accurate process model that simulates system behavior under normal conditions is required to run in parallel with the operating process. In reality, approximations in the models and drift in the system may result in the estimated state vector slowly deviating from the actual system values. To prevent this, an observer mechanism [6] is employed to make corrections to the estimated state vector. A critical issue with observers is the model adaptation rate, especially in case of incipient faults. If this rate is too fast, the model quickly adapts to changes in the system variables caused by the incipient faults, therefore, the generated nominal values do not indicate a deviation. The comparison of actual measurements to predicted nominal values of measured system variables leads to *fault detection*. To account for the effects of noise and measurement inaccuracies, based on design documentation, a margin of error is added to the nominal values to increase robustness and avoid false alarms [19]. When error thresholds are exceeded, the diagnosis system responds by setting corresponding alarms.

The monitoring stage plays a crucial role in successful fault detection, isolation, and refinement. Monitoring parameters such as *sampling rates* affect measurement interpretation, and, therefore, fault hypothesis generation and refinement. Depending on the monitoring implementation, certain faults may or may not be distinguishable from others, and this determines the overall diagnostic accuracy. A critical and related issue in FDI is sensor placement and measurement selection. This is an integral component of *diagnosability analysis*, i.e., choosing measurements that help isolate and differentiate among possible faults that may occur in the system [9], [14]. This paper develops an integrated framework for monitoring, prediction, and diagnosis from transients TRANSCEND, based on the architecture presented in Fig. 1.

II. Model Based Diagnosis

Our approach to diagnosis uses qualitative dependency relations between parameters and observed variables to generate hypothesized faults from observed deviations and to predict their future transient and steady state behavior.

A. Model Based Diagnosis System

In previous work, static models based on qualitative constraint equations [1] and signed directed graphs (SDG) [16] led to under constrained models that caused combinatorial problems in the diagnosis task. These system models did not incorporate dynamics, therefore, temporal feedback effects could not be dealt with, or had to be re-introduced on an ad hoc basis [15].

1) Modeling for Diagnosis: Generating successful models for diagnosis of continuous dynamic systems introduces a unique set of requirements.

- The models should describe both normal and faulty system behavior. The former provides the reference variable values for the monitoring task, and the latter forms the core for the prediction algorithm.
- The model should generate dynamic behavior under faulty conditions, so fault transients can be predicted by the model.
- The model should incorporate sufficient behavioral detail so deviations in observed variables can be mapped back to system components and parameters.
- When faults occur, the system may undergo a structural change. Analyzing structural changes is beyond the scope of this paper. However, they constitute an important category of failures, so it is important to not preclude them from the underlying framework. In parallel, we have been developing modeling techniques that combine discrete changes with continuous behavior analysis [11].

In addition, to constrain the inherently exponential search space for diagnosis, it is important that the model impose all relevant physical constraints on the search process. Also, given the limits of purely qualitative and purely quantitative schemes that have been discussed elsewhere [5], [6], [16], models that generate and use both qualitative and quantitative information are preferred to prevent loss of *a priori* information.

Analyzing the effects of abrupt faults is the key to successful fault isolation. Abrupt changes in the parameter values of energy storage elements may cause an abrupt change in some measured variables [10], [13]. To illustrate, assume that at time t_f, a rock falls into an open tank with capacity C, and an outflow resistance R for a connected outlet pipe (Fig. 2). The capacity of the tank decreases abruptly to C'. Since $p = \frac{q}{C}$, and q, the amount of liquid in the tank is conserved (assuming no overflow), the abrupt change in the capacitance value must reflect as an abrupt change in pressure, p, to $p' = \frac{C}{C'}p$. This does not have to be the case always. For example, an abrupt change in pipe resistance, may cause an abrupt change in outflow, but not an abrupt change in p.

2) Bond Graphs for Diagnosis: Bond graphs [17] provide a systematic framework for building consistent and well constrained models of dynamic physical systems across multiple

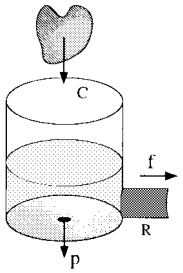


Fig. 2. Discontinuous change in tank capacitance.

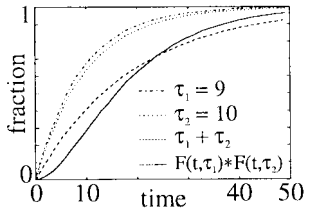
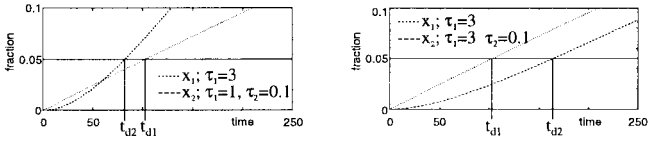


Fig. 3. Delay times of two first order systems (τ_1 and τ_2), their sum ($\tau_1 + \tau_2$), and the actual delay time of their combined effects ($F(t, \tau_1) * F(t, \tau_2)$).



(a) (b)

Fig. 4. Delay times for observing deviations.

domains. They include causality constraints that provide the mechanisms for effective and efficient diagnosis. An added advantage of bond graph representations is their direct applicability to qualitative processing, making them useful in situations where precise numerical information may not be available. Analytic system models derived from bond graphs are also amenable to quantitative simulation and analysis. Tangina *et al.* [21] derived analytic redundancy relations, and Linkens and Wang [7] compute local qualitative relations for fault isolation. Our work exploits the topological constraints of bond graph model for efficient diagnosis.

To extend bond graph modeling to component oriented diagnosis requires establishing correspondence between individual components and bond graph elements. In the bond graph framework, primitive elements, such as resistors and capacitors represent mechanisms which may not always be in one-to-one correspondences with individual system components [2]. An individual component may have multiple aspects represented in the bond graph. For example, a component such as a pipe may be represented in the bond graph by its build-up of flow momentum (I) and resistance to flow (R). Biswas and Yu [2] describe a compositional methodology for deriving bond graph models for diagnosis from a physical system description so that the bond graph elements directly correspond to system components and mechanisms under diagnosis scrutiny. The modeling methodology is further developed by Mosterman and Biswas [11], [13]. In our framework, a fault manifests as a deviation of a component parameter in the bond graph model.

Definition 2 (Fault): Faults are defined by model parameters that have deviated from their normal operating values.

B. Diagnosis from Transients

Abrupt faults like sudden blockages in pipes create transients in dynamic system behavior. This differs from a pipe that slowly accumulates dirt creating an incipient fault, which is more likely to cause a gradual drift in the system steady state behavior.

1) Characterizing Transients with Time Constants: Time constants play a key role in characterizing the dynamic behavior of physical systems. Faults cause instantaneous changes in some system variables. For other variables, energy storage elements acting as buffers introduce propagation delays and changes take longer to manifest. In general, variables with larger time constants take longer to produce observable changes when compared to variables with smaller time constants. If measurement snapshots are available from the system at rates that are faster than the smallest time constant, it becomes easier to track transients, and relate them to primary fault causes. In this work, without much discussion, this is assumed to be true.

Assumption 1 (Time Scale of Observation): Observations are sampled at rates that are faster than modeled system time constants in both normal and faulty operation.

Physical systems are inherently continuous, and hypothesized abrupt changes (e.g., the abrupt pressure change caused by the falling rock) actually occur on time scales smaller than the sampling rate for the observations. Therefore, they seem to manifest as discontinuous changes, but this is a sampling artifact attributed to the time scale of observation.

Definition 3 (Discontinuity): A change in a signal value that takes place on a time scale much smaller than the time scale of observations is classified as abrupt, and called a discontinuity.

Observed transients in system behavior may be affected by the combination of multiple time constants in subsystems that define the overall delay. The combined effects are a convolution rather than the sum of individual time constants [10]. As an illustration, Fig. 3 shows the step response of two first order systems with time constants τ_1 and τ_2, respectively. The combined effect of these systems is given by $F(t, \tau_1) * F(t, \tau_2)$ (convolution) whereas the sum of their individual delay times is shown by $\tau_1 + \tau_2$. Tracking of the measured values will produce significant error if the sum of the time constants is used instead of the convolution. This approach is further complicated by the fact that nominal time constants change when faults occur.

A qualitative framework mitigates this tracking problem to some extent but introduces problems in temporal ordering. Qualitatively, a measurement is considered *normal* if it is within a certain percentage (say 2%–5%) of its nominal value and deviant otherwise. Fig. 4(a) shows two variables affected by a fault, a first order effect, x_1, and a second order effect, x_2. The delay times, i.e., the time before these variables cross the error threshold are t_{d1} and t_{d2}, respectively. At times between

t_{d2} and t_{d1}, x_2 is reported deviant but x_1 is reported normal. Although x_2 embodies a second order effect with a zero value first-order derivative at the point of failure, it crosses the error threshold before a first order effect. This is contrary to expectations; a first order effect is expected to dominate (i.e., be much faster than) a second order effect. Fig. 4(b) compares two signals whose first order time constants are equal. The pure first order effect is faster than the signal that combines the first order effect with a second order effect.

This brings up an important issue when dealing with normal values and deviations from normal in a qualitative reasoning framework. A temporal ordering of first and higher order effects in studying deviations from normal is, in general, impossible unless the sensor system is wired and calibrated with extreme care to guarantee a temporal ordering in response times. Also, an observation being reported normal at a given time may actually be a slowly changing value that has not crossed the threshold, and, therefore, it should not be used to refute faults. This invariably produces contradictions in consistency based diagnoses.

In our approach, deviant observations are individually analyzed to generate sets of single fault hypotheses. Normal observations are not necessarily used to refute faults because it is hard to differentiate between a truly normal signal versus one that is changing slowly and will cross the normal threshold at some future point in time. Only in situations where discontinuities can be reliably detected can normal observations be used to refute faults that would cause a discontinuous change for that observation.

2) Feature Detection: Individual signal features are the prime discriminating factor between competing fault hypotheses. Signals can be noisy, therefore, prudence must be exercised in distilling information from them. Magnitude or zero order changes are measurable within a given error tolerance that is determined by the properties of the associated sensors. Slopes or first order derivatives can be reliably computed from measured signals in a qualitative framework [no change (0) and increasing or decreasing $(\pm)$] using standard filtering techniques. However, the measurement or derivation of higher order derivatives produces unreliable results [3]. Dedicated transducers, such as accelerometers, may be employed to measure second derivatives, but only for specific kinds of measurements. Therefore, our monitoring and feature detection subsystem focuses on making magnitude and slope measurements. Like magnitude, a slope that is currently within bounds and labeled normal (0) cannot be used to derive diagnostic conclusions because its value may change with time. Only when the measured slope deviates significantly from the expected value is this value directly used for fault isolation and refinement.

Specialized algorithms may be employed to derive other useful features from signals in a qualitative framework. For example, a simple discontinuous change detection mechanism can be based on observing that the magnitude and slope of an observed signal at the time point of failure have opposing signs. This discontinuity detection scheme has been successfully applied to systems in the hydraulics domain. Not all discontinuities take this form, and, therefore, the characteristic forms a necessary but not sufficient condition for discontinuity detection.

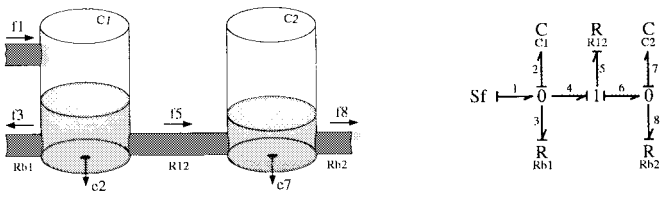


Fig. 5. Bi-tank system and its causally augmented bond graph model.

Another general characteristic of most physical systems is that dissipative effects eventually cause the system to return to a steady state. This translates to another feature that aids the fault isolation process. If it can be determined from the monitoring process that the eventual steady state will be above, below, or at the previous steady state value, one can distinguish between certain resistive and energy storage element faults.

Overall, our approach uses three features that take on the following values in our qualitative reasoning framework:

- magnitude:
  - —*low, high*;
  - —*discontinuity low, no discontinuous change, discontinuously high.*
- slope: *below normal, above normal*;
- steady state: *below, at, above original.*

III. Fault Hypotheses and Signatures

The general FDI methodology illustrated in Fig. 1 is implemented using bond graphs as the underlying modeling language. Dynamic characteristics of system behavior derived from the bond graph are represented as a *temporal causal graph.* Our algorithms for monitoring, fault isolation, and prediction are based on this representation. The fault analysis and refinement process continues till fault transients are masked by interactions or the system reaches a steady state. The goal is to uniquely identify the true fault using a combination of transient and steady state analysis.

A. The Temporal Causal Graph

The temporal causal graph is derived in two steps [12].

1) An extension of the SCAP algorithm [17], [22] is used to generate a graph that incorporates *cause-effect* relations among the power variables in the bond graph.
2) Component parameters and temporal information are added to individual causal edges to form the temporal causal graph. This adds temporal characteristics to the relations between variables.[2]

The temporal causal graph for the bi-tank system in Fig. 5 is shown in Fig. 6. The graphical structure represents effort and flow variables as vertices, and relations between the variables

[2]Note that the bond graph formalism presents one way to derive temporal causal graphs. Other modeling formalisms that support the physical modeling paradigm and allow for the generation of a temporal causal graph may be employed in its place.

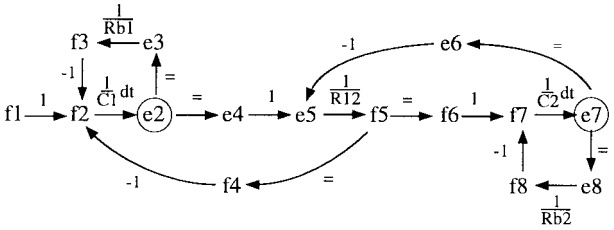


Fig. 6. Temporal causal graph of the bi-tank system.

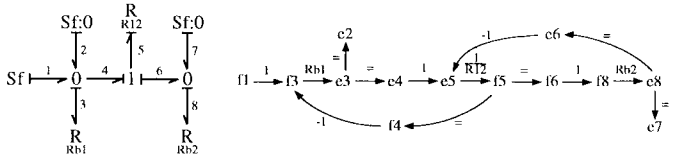


Fig. 7. Steady state bond graph of the bi-tank system and its corresponding causal graph.

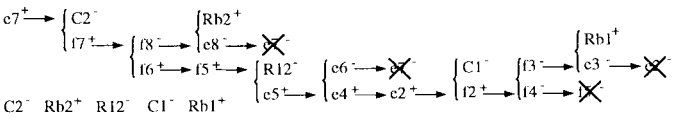


Fig. 8. Backward propagation given e_7^+ to find faults.

as directed edges. The relations can be attributed to junctions and system components. Junctions are of two types:

1) parallel or common effort (0-) junctions;
2) series or common flow (1-) junctions.

0-junctions require that the effort (e.g., pressure) values of all bonds incident on that junction be equal, and the sum of the flow (e.g., the fluid flow rate) values be zero. 1-junctions require that the flow values of all incident bonds be equal, and the effort values sum to zero. In the qualitative framework these relations impose labels -1, 1, and $=$ on graph edges. The $=$ implies that the junction constrains the two variable vertices associated with the edge to take on equal values, 1 implies a direct proportionality and -1 implies an inverse proportionality for the variables associated with the two incident vertices. An edge associated with a component represents the component's constituent relation. For example, the edge corresponding to a resistive element involved in an effort to flow relation is labeled $\frac{1}{R}$, and for a capacitor in integral causality the edge from flow to effort is labeled $\frac{1}{C}\,dt$.

Junctions, transformers, and resistors define instantaneous magnitude relations, whereas capacitors and inductors introduce magnitude and temporal effects on causal edges. In general, the temporal effects are *integrating*, and their associated rate of change is determined by the path that links an observed variable to the initial point where a deviation occurs. Natural feedback mechanisms in dynamic physical systems result in closed paths in the temporal causal graph (see Fig. 6). For loops with passive elements, these feedback mechanisms always have a negative gain [23] (e.g., the $f_7 \rightarrow e_7 \rightarrow e_8 \rightarrow f_8 \rightarrow f_7$ loop). Loops that include an integrating effect (e.g., $f_2 \rightarrow e_2 \rightarrow e_4 \rightarrow e_5 \rightarrow f_5 \rightarrow f_4 \rightarrow f_2$) are referred to as *state loops*.

Definition 4 (State Loop): A closed causal path with one and only one time-integrating effect is called a state loop. In previous diagnosis work, where temporal aspects of relations were not modeled explicitly, these ubiquitous negative feedback mechanisms caused difficulties in assigning deviation values in a consistent manner during the fault generation stage (see Section III-B). The problem was addressed by employing ad hoc criteria to break loops. In our work, this problem is easily addressed by exploiting the time delays in propagating signal values introduced by the integrating effects in state loops.

An added advantage of bond graph models is that they allow automatic derivation of the steady state model of the system. For the bi-tank system, both the tank capacities in steady state can be replaced by flow sources with value 0, since no change of stored energy takes place. The steady state bond graph and its resulting steady state causal graph are shown in Fig. 7. In steady state, causality assignments do not imply a temporal ordering, and a steady state graph represents a set of algebraic equations rather than differential equations. Therefore, causal links in the steady state graph have less meaning. Because the set of algebraic equations is invariant, parameter deviation effects do not change for different causality assignments.

B. Component Parameter Implication

For every recorded discrepancy between measurement and nominal value a backward propagation algorithm (Algorithm 1) is invoked on the temporal causal graph to implicate component parameters. Implicated component parameters are also labeled $-$ (below normal) and $+$ (above normal). The algorithm propagates observed deviant values backward along the directed edges of the temporal causal graph and consistent $-$ and $+$ deviation labels are assigned sequentially to vertices along the path if they do not have a previously assigned value. An example of its application is shown in Fig. 8 for a deviant pressure, e_7^+, in the right tank of the two tank system in Fig. 5. When e_7 is measured to be above its nominal value, backward propagation starts along $f_7 \overset{\frac{1}{C_2}dt}{\rightarrow} e_7^+$ and implicates C_2 as below normal (C_2^-) or f_7 as above normal (f_7^+). Backward propagation from f_7^+ along $f_6 \overset{1}{\rightarrow} f_7^+$ implies f_6^+, and the inverse relation on $f_8 \overset{-1}{\rightarrow} f_7^+$ implies f_8^-. Propagation along a path is terminated when a conflicting assignment is reached.

Backward propagation accounts for temporal effects by propagating deviant values along edges with instantaneous relations first. This ensures that no faults associated with higher order effects conflict with faults identified with lower order effects. An example is shown in Fig. 9. Backtracking along the path $e_1 \overset{1}{\leftarrow} e_2 \overset{dt}{\leftarrow} e_4 \overset{a}{\leftarrow} e_5$, hypothesizes a^+ as a fault consistent with the observation e_1^+, but the link $e_2 \overset{dt}{\leftarrow} e_4$ introduces a first order effect. However, the path $e_1 \overset{1}{\leftarrow} e_3 \overset{-1}{\leftarrow} e_4 \overset{a}{\leftarrow} e_5$ depicts a set of instantaneous relations that support the hypothesis, a^- implies e_1^+. At the point of failure, the instantaneous $a^- \rightarrow e_1^+$ effect dominates the first order $a^+ \rightarrow e_1^+$ effect. When analyzing an individual

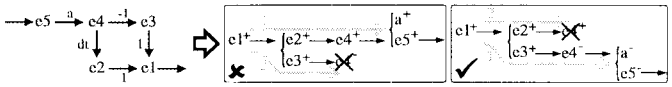


Fig. 9. Instantaneous edges propagate first.

signal's transient behavior, it is clear that its lower order effects manifest before its higher order effects. Therefore, the backward propagation algorithm is designed to propagate in a depth-first manner in increasing order of time-derivatives. All component parameters along a propagation path are possible faults. As discussed earlier, observed normal measurements do not terminate the backward propagation process. The result of backward propagation is a set of hypothesized single faults that are consistent with the reported deviant observations.

Algorithm 1 Identify Possible Faults

$f_{list} = \emptyset; v_{list} = \emptyset$
for all reported discrepant measurements **do**
  add vertex corresponding to deviant measurement to v_{list}
  and mark vertex with qualitative deviation value
  while v_{list} is not empty **do**
    $v_{current} \leftarrow$ the first vertex in v_{list} (and delete $v_{current}$ from v_{list})
    while $v_{current}$ has unmarked ancestors **do**
      if ancestor relation includes a parameter **then**
        add the parameter with consistent label to the f_{list}
      end if
      if ancestor vertex is unmarked **then**
        ancestor value $\leftarrow$ **new_value**(current value, relation)
        if relation is instantaneous **then**
          add the ancestor vertex to the beginning of v_{list}
        else
          add the ancestor vertex to the end of v_{list}
        end if
      end if
    end while
  end while
  $v_{list} = \emptyset$
end for

C. Prediction

Once faults are hypothesized, prediction and refinement schemes are employed to converge on the true fault. A more complete prediction module may be required to handle model changes when faults cause structural changes in the system. We assume faults do not cause changes in system configuration, and the system model remains valid even after faults occur in the system.

Assumption 2 (No Structural Changes): Faults do not cause the system model to undergo configuration changes. The prediction module uses the system model to compute the dynamic, transient, behavior of the observed variables and the eventual steady state behavior of the system under the fault conditions. Future behavior is expressed in qualitative terms: magnitude (zeroth order time-derivative), slope (first order time-derivative) and higher order effects.

Definition 5 (Signature): The prediction of zeroth, first, and higher order time-derivative effects of a system variable as qualitative values: below normal (*low*), *normal*, and above normal (*high*) in response to a fault is called its signature.

Algorithm 2 Predict Future Behavior for a Fault

add initial vertex, i.e., immediate consequence of the fault to list v_{list}
mark vertex 0^{th} order derivative with qualitative value
while v_{list} is not empty **do**
  $v_{current} \leftarrow$ the first vertex in v_{list}
  while $v_{current}$ has successors `not determined to sufficient order` **do**
    if successor relation includes a time integral effect **then** increase current derivative order
    end if
    if derivative order $\leq$ maximum order **then**
      if successor derivative is `no_mark` **then**
        successor derivative value $\leftarrow$ **new_value**(current value, relation)
      else if successor derivative has opposite value of current **then**
        successor derivative value $\leftarrow$ `conflict`
      end if
      add the successor to end of v_{list}
    end if
  end while
end while
for all vertex derivatives **do**
  if value = `no_mark` and any higher order derivative $\neq$ `no_mark` **then**
    replace `no_mark` with `normal`
  end if
  if value = `conflict` **then**
    replace `conflict` with `no_mark`
  end if
end for

The method for predicting future system behavior is presented as Algorithm 2. The algorithm propagates the effects of a hypothesized fault to establish a qualitative value for all measured system variables. Forward propagation along temporal edges implies an integral effect, therefore, the cause variable affects the derivative of the effect variable. All deviation propagations start off as zeroth order effects, i.e., as magnitude changes. When an integrating edge is traversed, the magnitude change becomes a first order change, i.e., the first derivative of the affected quantity changes. This is illustrated

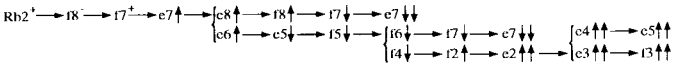


Fig. 10. Forward propagation forimplicated component R_{b2}^{+} to establish its signature.

by an $\uparrow$ ($\downarrow$) in the propagation example in Fig. 10. Similarly, a first order change propagating across an integrating edge produces a second order change ($\uparrow\uparrow$ ($\downarrow\downarrow$)), i.e., the second derivative of the affected variable changes. Second order changes propagate to third order changes, and so on. The forward propagation algorithm operates breadth-first along the temporal causal graph.

The forward propagation algorithm terminates when a signature of *sufficient* order is generated. A *complete* signature for an observation contains derivatives specified to its sufficient order. The sufficient order of a signature depends on the set of chosen measurement variables and the desired level of *diagnosability* for the system.

Definition 6 (Diagnosability): Diagnosability is a function of the number of possible faults that can be uniquely identified by a fault isolation system.

A fault detection scheme is *completely diagnosable* for a given set of measurements if it can isolate all possible single faults with the set of measurements.

Definition 7 (Complete Diagnosability): A fault isolation system is completely diagnosable if it can uniquely isolate all possible hypothesized faults.

Diagnosability depends on the selected observation set and the chosen order of their signatures [14]. In theory, consideration of higher order variable effects is likely to results in greater diagnosability. Therefore, the same diagnosability can be achieved with a smaller number of total observations but considering higher order signatures, or using a larger number of observations with lower order signatures. In practice, using signatures of lower order has advantages. Higher order effects take longer to manifest, and fault patterns take longer to establish after failure occurs. During this time, feedback effects in the system may be superimposed on initial fault behavior and change the nature of the patterns. This problem is compounded even further when cascading faults occur.

The steady state causal graph derived from the bond graph model of the system determines the final steady state value for each observed variable under the faulty conditions. The predicted steady state value for each observed variable, i.e., *below the original, at the original*, or *above the original* steady state, is attached to the signature and used in the monitoring stage.

IV. Monitoring

The monitoring module compares predicted signatures of the hypothesized faults to actual measurements as they change dynamically. A number of issues of practical importance, related to the quality and characteristics of the measurements, are incorporated into the monitoring scheme so dynamic effects can be realistically measured using local mechanisms.

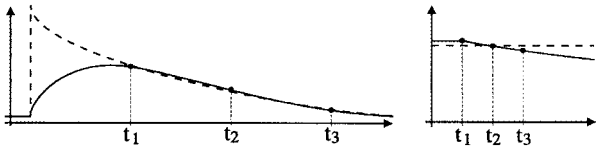


Fig. 11. Signal interpretation.

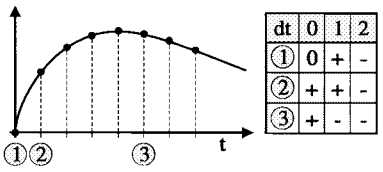


Fig. 12. Progressive monitoring.

A. Sensitivity to the Time Step

The monitoring time step is critical to the success of the overall fault isolation scheme. The step size depends on the different rates of response that the system exhibits. Too large a time step may produce incorrect inferences as shown for the signal (solid curve) at the left in Fig. 11. A large monitoring time step ($>t_1$) gives the appearance that this signal undergoes a discontinuous change (dashed curve). Decreasing the time step helps in differentiating discontinuities (abrupt changes) from continuous effects but if the time step is too small, it appears that the signal does not change for a large number of steps (see plot on right of Fig. 11). Too small a time step may result in lack of sensitivity to changes and unnecessary computational expense on the analysis task. If the variable corresponding to a slowly changing measurement is prematurely reported to be normal, or to have reached a new steady state value, this may result in elimination of true faults. To mitigate this problem, fault refutation based on a given observation is only invoked after an initial deviation is detected. As discussed earlier, the sampling rate also determines whether the effect of a fault is observed to be discontinuous.

B. Progressive Monitoring

Transient characteristics at the time of failure tend to change over time as other phenomena in the system affect the measured variables. The signatures for a candidate fault can change dynamically. For example, a fault in the system may have no effect on the initial magnitude (0th order value) of a variable, but it may affect its 1st derivative, predicting that it will be above normal. Therefore, immediately after the fault occurs the variable value will be observed to be normal (its deviation is within the 2–5% threshold), but as time progresses, the derivative effect will cause the variable value to go above normal. Fig. 12 depicts time stamps marked 1, 2, and 3, where a lower order predicted effect is replaced by a higher order effect. The notion of employing higher order derivatives in analyzing measured variables during the monitoring process is referred to as *progressive monitoring*. When as observed variable does not match a predicted normal value, the comparison is successively extended to predicted higher order derivatives in the variables signature. If the higher

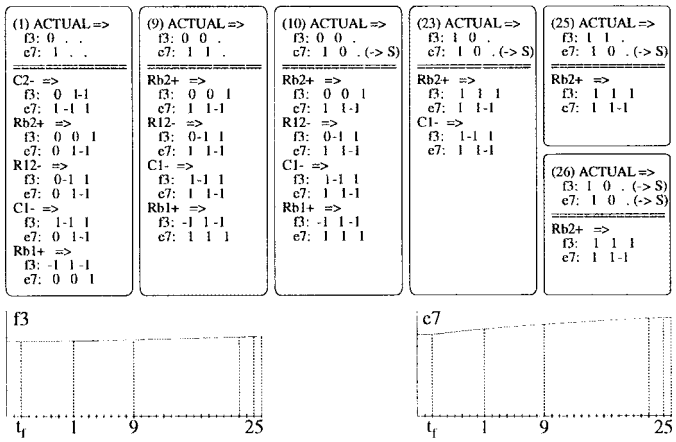


Fig. 13. Progressive monitoring forfault R_{b2}^{+}.

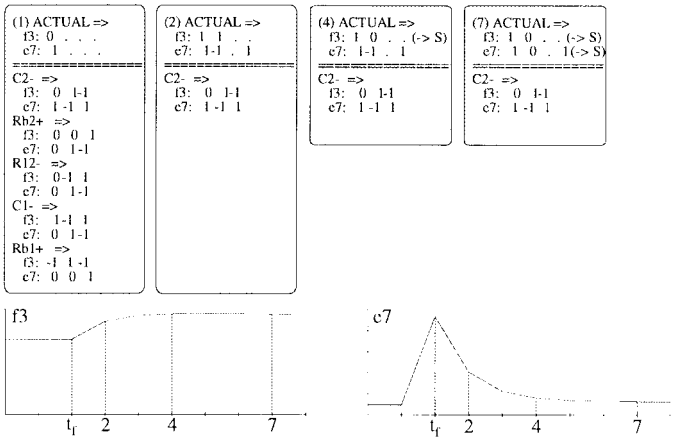


Fig. 14. Results of the diagnosis system with C_2^{-} faulty and discontinuity detection is used.

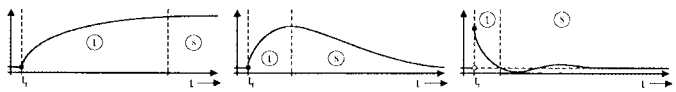


Fig. 15. Typical signal transients in physical systems that exhibit different qualitative behavior over time.

order derivatives match the observed value, the hypothesized fault is considered plausible, otherwise it is rejected.

To illustrate, Fig. 13 shows the predicted and monitored behavior for a sudden increase in outflow resistance R_{b2} in the bi-tank system in Fig. 5, where -1, 0, 1 maps onto *low, normal, high* and a period indicates that the value is *unknown*. The two observed variables are the outflow of the left tank, f_3, and the pressure in the right tank, e_7. Not all monitoring output is shown; the boxes depict the monitored values at time steps where the set of hypothesized faults changes or where the tracking of an observation's transient behavior is terminated. The actual observations and the newly inferred set of possible faults and their signatures are listed. The values on the top of each box represent the measured signal magnitude (zeroth order), slope (first derivative), and second derivative[3] expressed in qualitative terms. Below the reported measurements are the predicted signatures of the measured variables for each hypothesized fault. Consider fault R_{b2}^{+} and measurement f_3 in Fig. 13. At step 9, the reported value for f_3 is still normal (its value has not exceeded the error threshold), and this agrees with the signature 0, 0, 1 for R_{b2}^{+}. At step 23, the reported value for f_3 is 1, 0 (magnitude above normal), which no longer appears to be consistent with fault R_{b2}^{+}'s signature. However, when progressive monitoring is applied, the second derivative, which is positive, makes an impact on both the first derivative and magnitude of the signal, and the prediction for R_{b2}^{+} is changed to 1, 1, 1. Updating the prediction in this manner keeps the signature consistent with the observation, and R_{b2}^{+} is still a viable fault hypothesis. Hypothesized faults are dropped if their signatures do not match observations. Note that in step 23 the slope for f_3 is reported to be 0 whereas the magnitude deviates. This is an artifact of our implementation as the deviation in a slope is computed using the first set of observations after an initial magnitude deviation is detected.

Fig. 14 illustrates progressive monitoring with discontinuity detection (see Section II-B). The change in f_3 and e_7 when the fault C_2^{-} occurs in the bi-tank system is listed. A fourth field is added to the actual observation; a value of 1 in this field implies a positive discontinuous deviation occurred at the time of failure. Matching an initial discontinuous change produces a unique fault after the second time step. The discontinuous change observed for measurement e_7 at step 2 implicates C_2^{-} and the other hypothesized faults are eliminated. The flow f_3 is observed to have a positive deviation and positive slope $(1, 1)$ as opposed to $(1, -1)$ for e_7. Therefore, the change in f_3 is not labeled discontinuous by our criteria. The discontinuity detection criteria is a necessary but not sufficient condition.

The diagnosis engine can correctly detect and isolate all single fault parameter deviations if the pressure in one tank and the outflow from the other or the flow between them were measured and first order signatures are used. In this case, discontinuity detection is not required but steady state detection is. If steady state detection is not feasible, three observations and discontinuity detection have to be used, or a second order signature without discontinuity detection can be employed. The task of measurement selection to achieve complete diagnosability is discussed in greater detail elsewhere [14]. Detailed results for two-tank and three-tank systems are presented in [12].

C. Temporal Behavior

Two distinct characteristics of signals in response to fault disturbances, transients and steady state, carry the most distinctive discriminative information for diagnosis. For monitoring it is important to know when, after a time of failure t_f, the transient detection phase terminates, and the system moves into the steady state mode, requiring steady state detection to be activated.

Palowitch [16] reports that signals may exhibit a *compensatory* [Fig. 15(a)] or an *inverse response* [Fig. 15(b)].

[3] As discussed earlier, second and higher order derivatives are not measured. This slot is retained to make it easier to match actual and predicted observations.

A compensatory response exhibits a decreasing slope and gradually moves toward a new steady state value. For an inverse response, after an initial increase or decrease, the signal may reverse direction. An additional phenomenon resulting from abrupt faults can be categorized as a *reverse response* [Fig. 15(c)]. A reverse response occurs if a discontinuous signal overshoots, and, consequently, its qualitative magnitude reverses sign (i.e., goes from above normal to below normal or *vice versa*).

In the qualitative analysis framework, the transition to steady state analysis is detected from an initial magnitude deviation by noting the following.

- For a compensatory response, the slope eventually becomes 0.
- For an inverse response no discontinuous change of magnitude is associated with t_f. The switch from transient to steady state detection occurs when the magnitude and slope deviations take on opposing signs. Eventually the slope may become 0.
- For a reverse response the signal has a discontinuous initial magnitude deviation with sign that is opposite of the current magnitude deviation. The switch to steady state detection occurs when the magnitude changes sign.

When any of these situations are detected, transient verification for that particular signal is suspended (stage t in Fig. 15), and steady state detection is activated (stage s in Fig. 15). After a period of time, some signals may be processed in the transient mode, whereas others are processed in the steady state mode. Steady state is detected when a first order derivative becomes 0 for a sufficient period of time. The sufficient period of time is usually based on design information. The $\rightarrow$ S in Fig. 13 illustrates that transient detection was suspended for e_7 from time step 10. At this point in time, steady state detection is activated for this signal only. At step 26 transient detection for f_3 is suspended and steady state detection is initiated. Both these are examples of a compensatory response. However, the difficulty in detecting the final steady state value results in it not being used as a verification mechanism here, and, the diagnosis process ends at time step 26. In Fig. 14, the diagnosis process terminates at step 7. As part of future research, more sophisticated steady state detection techniques will be investigated.

D. Summary

Monitoring plays a key role in the robustness of the fault analysis scheme. The following issues summarize the monitoring and measurement selection process.

1) Only deviating signals play a role in transient fault analysis. This circumvents the problem of insensitivity to small time steps.
2) During this transient monitoring stage a progressive monitoring scheme defines the dynamic characteristics of the initial fault transients.
3) After a period of time, signal behavior may deviate significantly from transient behavior at the time of failure (e.g., it may reverse its slope). In this situation, the transient prediction and verification process is suspended, and steady state analysis is activated. This is based on the three characteristic qualitative signal behaviors discussed earlier in this section. Suspension of transient analysis and steady state detection are non trivial tasks in the monitoring and fault isolation scheme.
4) An off-line measurement selection algorithm [14] identifies the sufficient order of predictions for fault isolation to achieve a degree of diagnosability.

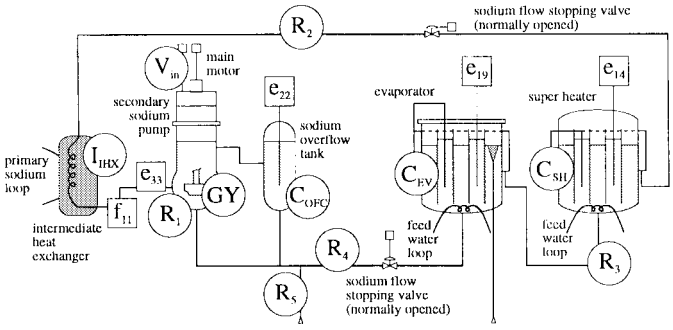


Fig. 16. Secondary sodium cooling loop.

V. Liquid Sodium Cooling System

The scalability of our FDI methodology was tested by conducting experiments on the simulation model of the secondary liquid sodium cooling loop in a fast breeder nuclear reactor. The need for a qualitative approach in this system is motivated by the fact that it is modeled as a nonlinear sixth order, system. This makes it hard to develop accurate numeric models for generating system behavior in different modes. Moreover, the precision of the sodium flow sensors used in the system is limited and hardware redundancy is difficult to achieve because of the expense involved in adding flow sensors.

A. Secondary Sodium Cooling Loop

In a nuclear reactor, heat from the reactor core is transported to the turbine by a primary and secondary cooling system. The primary cooling sub-system connects directly to the reactor and transfers heat to the secondary cooling sub-system which then transfers heat carried by the liquid sodium to the steam in the generator (Fig. 16). Heat transfer from the primary cooling loop to the liquid sodium in the secondary loop happens through an intermediate heat exchanger. The heated sodium is then pumped through two stages: the super heater and the evaporator vessel, both of which heat up the water and steam in the steam–water loop that then drives the turbine.

1) Bond Graph Model: The model used for diagnosis applies energy and mass balance of the system in the hydraulics domain combined with the mechanical characteristics of the main motor and pump. The bond graph that captures system behavior in these domains is a nonlinear, sixth order model (Fig. 17). The main motor driver (Fig. 18) is a synchronous ac motor. As a simplification the electrical subsystem is not modeled. The electrical part of the motor system can be represented as a source of mechanical energy with a given torque/angular velocity characteristic. The inertia of the rotor and the mass of the transmission gear is modeled by m_1, and

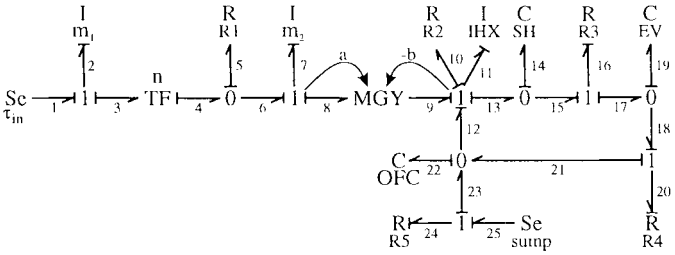


Fig. 17. Bond graph model for the secondary sodium cooling system.

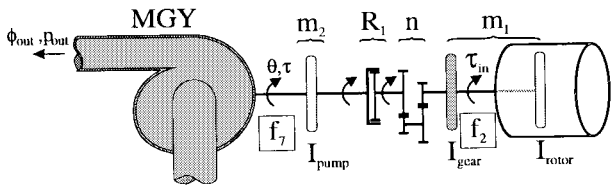


Fig. 18. Synchronous ac motor that drives a pump.

the transmission ratio between motor and pump by n. Pump losses in the fluid connection between the motor and pump are modeled by a dissipation element, R_1, and the pump inertia is represented as m_2. The model of a centrifugal pump can be derived using conservation of power and momentum [23]. Given that the amount of mass moved by the pump depends on the total area of its veins, a, minus the effective loss in moved mass due to the curvature of the veins, b, this yields the constituent relations $\tau = (a\theta - b\phi_{\text{out}})\phi_{\text{out}}$ and $p_{\text{out}} = (a\theta - b\phi_{\text{out}})\theta$. This describes a modulated gyrator with modulus $a\theta - b\phi_{\text{out}}$, where $b = 0$ if the pump veins are not curved.

The hydraulics of the sodium loop are modeled by a closed power loop (Fig. 17). The coil in the intermediate heat exchanger accounts for flow momentum build-up, represented by a fluid inertia, I_{IHX}. The piping from the main pump through the heat exchanger to the evaporator vessel is represented by resistance R_2. The two sodium vessels are modeled by capacitances, C_{EV} and C_{SH} and the connecting pipe by its resistance, R_3. An overflow column, C_{OFC}, maintains a desired sodium level in the main motor, and the piping between the evaporator and this column is represented by resistance R_4. This storage facility is connected to a sump, S_e, by a pipe with resistance, R_5.

Solving the algebraic equations in the steady state model (i.e., all C elements in Fig. 17 are replaced by $S_f : 0$ and all I elements by $S_e : 0$) results in a third order equation because of the quadratic modulus of the gyrator. A closed form symbolic solution was derived using Mathematica. This solution has one real root that represents the steady state solution, and symbolically provides the values for nominal operation.

2) The Temporal Causal Graph: The temporal causal graph (Fig. 19) of the system is derived from the bond graph in Fig. 17. Because of its nonlinear character, the *MGY* requires more detailed analysis. The derivation of the causal relations of the modulated gyrator is shown in Fig. 20. First it is observed that the modulation factor $g = af_8 - bf_9$ is directly proportional to f_8 and inversely proportional to f_9. The dependency of g on f_8 and $-f_9$ can be explicitly

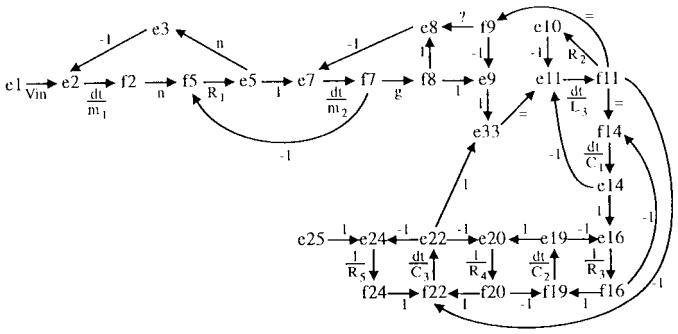


Fig. 19. Temporal causal graph of dynamic behavior.

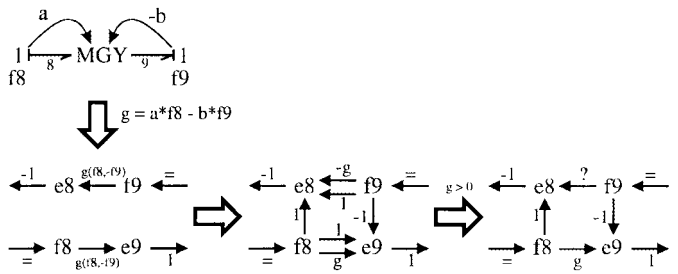


Fig. 20. Temporal causal graph of a modulated gyrator.

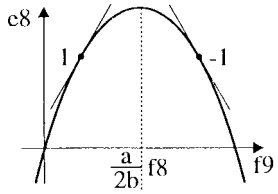


Fig. 21. Detailed sensitivity analysis of $\partial e_8/\partial f_9$.

modeled by edges between these variables and the affected variables. The bond graph indicates that f_8 and f_9 affect e_8 and e_9. The corresponding edges are added to the causal graph (Fig. 20). The added influences on e_8 result in ambiguity. This is revealed by studying the relation between e_8 and f_9. From the bond graph $e_8 = (af_8 - bf_9)f_9 = af_8f_9 - bf_9^2$. The plot in Fig. 21 reveals that the e_8 to f_9 link can have a positive or negative value depending on the values of f_8 and f_9. From nominal steady state values, the link sign can be precomputed. However, once a fault occurs, changes in the values of f_8 and f_9 may cause a change in the sign. Since the exact values of the two variables are not known at monitoring time, the sign on the link may or may not reverse. The reversal occurs only when e_8 is predicted to be high based on the proportional influence (-1 or 1). Since a predicted decrease in e_8 is unambiguous it is propagated, but a predicted increase in e_8 is propagated as *unknown*. The two pump parameters a and b are represented by one positive parameter, g, that is linked to pump fault.

B. Simulation Results

The numerical simulation model for the secondary cooling loop utilizes the forward Euler integration, $x_{k+1} = \dot{x}\Delta t + x_k$.

TABLE I
SECONDARY SODIUM COOLING LOOP PARAMETER VALUES

R_1	10	C_{SH}	20	m_1	0.1	n	0.25
R_2	1	C_{EV}	20	m_2	0.5	a	1
R_3	0.143	C_{OFC}	1.6	I_{IHX}	1	b	0.1
R_4	0.232						
R_5	1	τ_{in}	7	sump	2	EV_{max}	2.2

To ensure stability, the numerical time step Δt was chosen to be $< \frac{\tau_{\min}}{3}$, where $\tau_{\min}$ is the smallest time constant of the model [23].

From system specification documents and by consulting domain experts, the parameter values listed in Table I were chosen. Those values do not represent actual system parameter values, but their relative magnitudes are such that the generated qualitative characteristics of the behavior are correct. The $\mathrm{EV}_{\max}$ parameter indicates the maximum level of the liquid sodium in the evaporator vessel. The overflow mechanism was modeled but not included in the temporal causal graph to avoid model configuration changes. The simulation used a numerical time step of $\Delta t = 0.001$, which produced numerically stable simulations in both normal and failure situations. Thus $\tau_{\min} < 0.003$ for the model, which is equivalent to a minimal time constant in the order of minutes for the actual system. This was also chosen as the monitoring time step, which results in a sampling rate of about 20 s for the actual system.

Failure was simulated in the system by changing the model parameters by a factor of five. Conservation of state [13] was applied when capacitance and inductance failures occurred. Keeping the stored momentum or the amount of liquid constant resulted in an abrupt change of angular velocity/flow or pressure, respectively. Simulation was stopped when either the transients of all observations were detected or 3913 samples had been processed.[4]

The quality of the results depended on the parameter differences in the model and unmodeled configuration changes. For detection of high and low values for signals, a qualitative margin of error of 2% was used to avoid spurious deviations due to noise.

Table II summarizes the results. Columns 1 and 4 are the introduced faults, column 2 and 5 list the faults reported by TRANSCEND and columns 3 and 6 indicate the number of measurement samples required to arrive at the result. Three faults, R_3^-, R_4^+, and C_{EV}^-, were not accurately detected or isolated. Because of the overflow mechanism in the evaporator vessel, a decrease in capacity, C_{EV}^-, does not result in an increase in level and this is not detected. To detect this failure, flow of sodium through the overflow mechanism has to be monitored. The two other faults, R_3^- and R_4^+, were detected but not correctly isolated, again because the overflow mechanism was not modeled in the temporal causal graph. If this phenomenon is included by tagging a predicted value *unknown* instead of *high* when it would have predicted an evaporator level that is *high*, the faults would be accurately isolated as indicated by the entries in parentheses in Table II. Not all faults can be uniquely isolated ($\{n, R_1\}$) because of the lack of required measurements, or certain predicted deviations are too small to be observed.

[4]This number is derived from the time it takes a signal with time constant 1 to reach its steady state value within 2%.

TABLE II
FAULT DETECTION FOR $\{f_2, f_7, f_{11}, e_{14}, e_{19}, e_{22}, e_{33}\}$ WITH $\Delta t = 0.001$, order = 3, $q_{\mathrm{margin}} = 2\%$

Fault	Diagnosis	Samples	Fault	Diagnosis	Samples
R_1^+	n^+, R_1^+	58	R_1^-	n^-, R_1^-	43
R_2^+	R_2^+	27	R_2^-	R_2^-	46
R_3^+	R_3^+	1255	R_3^-	$\emptyset$ (R_3^-)	699 (699)
R_4^+	R_5^+ (R_4^+, R_5^+)	3429 (378)	R_4^-	R_4^-	43
R_5^+	$n^+, R_1^+, R_2^-, R_3^+, R_4^+, R_5^+$	2	R_5^-	R_3^-, R_4^-, R_5^-	687
C_{SH}^+	C_{SH}^+	73	C_{SH}^-	C_{SH}^-	16
C_{EV}^+	C_{EV}^+	45	C_{EV}^-	-	-
C_{OFC}^+	C_{OFC}^+	9	C_{OFC}^-	C_{OFC}^-	3
m_1^+	m_1^+	6	m_1^-	m_1^-	2
m_2^+	m_2^+	2	m_2^-	m_2^-	2
I_{IHX}^+	I_{IHX}^+	16	I_{IHX}^-	I_{IHX}^-	2

VI. CONCLUSIONS

This paper presents an effective mechanism for fault isolation in complex dynamic systems by analysis of qualitative transients and progressive monitoring of the evolving behavior of the system after initial fault occurrences. The work makes a number of important contributions:

1) use of the bond graph language to develop a systematic framework for dynamic and steady state analysis of physical systems;
2) use of qualitative signatures defined by higher order derivatives for tracking system behavior based on hypothesized faults;
3) progressive monitoring scheme for comparing evolving temporal system behavior to the signatures for fault refinement.

A number of experiments with two-tank and three-tank systems have produced excellent results. To demonstrate the value of the system in more realistic situations, we have applied it to a complex, sixth order model of a secondary sodium cooling loop system for a fast breeder reactor. Results obtained are encouraging, and the difficulties encountered are not an issue of scalability, but more the ability to model complex nonlinearities, the time-scales of different subsystems, and processing the effect of structural changes in the system.

Currently we are focusing on improving the analysis of the dynamic transients in the fault isolation mechanisms by incorporating order of magnitude relations of the temporal effects of integrating edges and developing more sophisticated discontinuity detection algorithms. We are developing systematic methods for handling structural faults like leaking pipes that cause changes in system configuration. We are also designing and implementing an environment for monitoring and analyzing real data from an operating automobile engine. This presents interesting challenges for developing signal interpretation techniques that are robust to noise, and the development of real time monitoring, prediction, and fault isolation algorithms.

ACKNOWLEDGMENT

The authors acknowledge the help of S. Yoshikawa and Dr. T. Washio in constructing the secondary sodium cooling loop models.

REFERENCES

[1] G. Biswas, R. Kapadia, and X. W. Yu, "Combined qualitative quantitative steady state diagnosis of continuous-valued systems," *IEEE Trans. Syst., Man, Cybern. A*, vol. 27, pp. 167–185, Mar. 1997.

[2] G. Biswas and X. Yu, "A formal modeling scheme for continuous systems: Focus on diagnosis," in *Proc. IJCAI-93*, Chambery, France, Aug. 1993, pp. 1474–1479.

[3] M. J. Chantler, S. Daus, T. Vikatos, and G. M. Coghill, "The use of quantitative dynamic models and dependency recording for diagnosis," in *Proc. 7th Int. Workshop Principles Diagnosis*, Val Morin, P.Q., Canada, Oct. 1996, pp. 59–68.

[4] R. N. Clark, P. M. Frank, and R. J. Patton, "Introduction," in *Fault Diagnosis in Dynamic Systems: Theory and Applications*, R. Patton, P. Frank, and R. Clark, Eds. Englewood Cliffs, NJ: Prentice-Hall, 1989, ch. 1, pp. 1–19.

[5] P. Frank, "Fault diagnosis: A survey and some new results," *Automatica*, vol. 26, no. 3, pp. 459–474, 1990.

[6] R. Isermann, "A review on detection and diagnosis illustrate that process faults can be detected when based on the estimation of unmeasurable process parameters and state variables," *Automatica*, vol. 20, no. 4, pp. 387–404, 1989.

[7] D. A. Linkens and H. Wang, "Qualitative bond graph reasoning in control engineering: Fault diagnosis," *Int. Conf. Bond Graph Modeling Simulation*, Las Vegas, NV, 1995, pp. 189–194.

[8] J. Mauss, "Diagnosis by algebraic modeling and fault-tree induction," in *Proc. 6th Int. Workshop Principles Diagnosis*, Goslar, Germany, Oct. 1995, pp. 73–80.

[9] A. Misra, J. Sztipanovits, A. Underbrink, R. Carnes, and B. Purves, "Diagnosability of dynamical systems," in *Proc. 3rd Int. Workshop Principles Diagnosis*, Rosario, WA, Oct. 1992, pp. 239–244.

[10] P. J. Mosterman, "Hybrid dynamic sysytems: A hybrid bond graph modeling paradigm and its application in diagnosis," Ph.D. dissertation, Vanderbilt Univ., Nashville, TN, 1997.

[11] P. J. Mosterman and G. Biswas, "Formal specifications for hybrid dynamical systems," in *Proc. IJCAI-97*, Nagoya, Japan, Aug. 1997, pp. 568–573.

[12] ———, "An integrated architecture for model-based diagnosis," in *Proc. 7th Int. Workshop Principles Diagnosis*, Val Morin, P.Q., Canada, Oct. 1996, pp. 167–174.

[13] P. J. Mosterman and G. Biswas, "A theory of discontinuities in dynamic physical systems," *J. Franklin Inst.*, vol. 335, no. 6, pp. 401–439, 1998.

[14] P. J. Mosterman, G. Biswas, and N. Sriram, "Measurement selection and diagnosability of complex dynamic systems," in *Proc. 8th Int. Workshop Principles Diagnosis*, Oct. 1997.

[15] O. O. Oyeleye, F. E. Finch, and M. A. Kramer, "Qualitative modeling and fault diagnosis of dynamic process by MIDAS," in *Readings in Model-Based Diagnosis*, W. Hamscher, L. Console, and J. de Kleer, Eds. San Mateo, CA: Morgan Kaufmann, 1992, pp. 249–254.

[16] B. L. Palowitch, "Fault diagnosis of process plants using causal models," Ph.D. dissertation, Mass. Inst. Technol., Cambridge, Aug. 1987.

[17] R. C. Rosenberg and D. Karnopp, *Introduction to Physical Systems Dynamics*. New York: McGraw-Hill, 1983.

[18] M. Sampath, R. Sengupta, S. Lafortune, K. Sinnamohideen, and D. C. Teneketzis, "Failure diagnosis using discrete-event models," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1555–1575, Sept. 1995.

[19] H. Schneider and P. M. Frank, "Observer-based supervision and fault detection in robots using nonlinear and fuzzy logic residual evaluation," *IEEE Trans. Contr. Syst. Technol.*, vol. 4, pp. 274–282, May 1996.

[20] J. L. Stein, "Modeling and state estimator design issues for model-based monitoring systems," *Trans. ASME*, vol. 115, pp. 318–327, June 1993.

[21] M. Tangina, J. P. Cassar, G. Dauphin-Tanguy, and M. Staroswiecki, "Monitoring of systems modeled by bond graphs," in *Int. Conf. Bond Graph Modeling Simulation*, Las Vegas, NV, 1995, pp. 275–280.

[22] J. van Dijk, "On the role of bond graph causality in modeling mechatronic systems," Ph.D. dissertation, Univ. Twente, The Netherlands, 1994.

[23] J. J. van Dixhoorn and P. C. Breedveld, *Technische Systeemleer*, 4th ed. The Netherlands: Univ. Twente, Feb. 1985.

Pieter J. Mosterman was born in The Netherlands on March 16, 1967. He received the B.Sc. and M.Sc. degrees from the University of Twente, Twente, The Netherlands, in 1987 and 1991, respectively. In 1997, he received the Ph.D. degree in electrical and computer engineering from Vanderbilt University, Nashville, TN.

Since 1997, he has been a Research Associate with the Control Design Engineering Group, Institute of Robotics and System Dynamics, German Aerospace Center. As part of his thesis research, he investigated principles that govern discontinuous behavior in physical systems. The thermodynamic properties of bond graphs were exploited to extend the modeling formalism to hybrid bond graphs, which was further developed into the rigorous mathematical representation to facilitate simulation and analysis. In the area of model based diagnosis, he developed algorithms for monitoring, prediction, and isolation of abrupt faults in physical systems. Currently, he is investigating the requirements for modeling complex dynamic systems, which involves integrating sophisticated formalisms for modeling and simulating both continuous time and discrete event phenomena.

Gautam Biswas (S'78–M'82–SM'91) received the B.Tech. degree in electrical engineering from the Indian Institute of Technology, Bombay, India, in 1977, and M.S. and Ph.D. degrees in computer science from Michigan State University, East Lansing, in 1979 and 1983, respectively.

He is an Associate Professor of computer science and engineering and management of technology at Vanderbilt University, Nashville, TN. He conducts research in artificial intelligence with primary interests in modeling and analysis of complex systems and their applications to diagnosis, design, and control. He is also involved in developing simulation based environments for learning and instruction. His research is currently supported by NSF, ONR, PNC Japan, and HP Research Labs. He has published in a number of journals and contributed book chapters.

Dr. Biswas has served on the program committees of a number of conferences. He was co-chair of the 1996 Principles of Diagnosis Workshop, and on the Senior Program committee for AAAI-97 and AAAI-98. He is a senior member of the IEEE Computer Society, ACM, AAAI, and the Sigma Xi Research Society.