N93-11923

A Model-based Reasoning Approach to Sensor Placement for Monitorability

Steve Chien, Richard Doyle, and Luiz Homem de Mello

Artificial Intelligence Group, Jet Propulsion Laboratory, M/S 301-490,
California Institute of Technology
4800 Oak Grove Drive, Pasadena, CA 91109-8099

Abstract

This paper presents an approach to evaluating sensor placements to maximize monitorability of the target system while minimizing the number of sensors. The approach uses a model of the monitored system to score potential sensor placements on the basis of four monitorability criteria. The scores can then be analyzed to produce a recommended sensor set. An example from our NASA application domain is used to illustrate our model-based approach to sensor placement.

Introduction

Sensor placement is the task of determining a set of sensors which allows the most accurate determination of the overall state of a monitored system while minimizing sensor power consumption, cost, computing power requirements, and weight. Reducing these quantities is particularly important in space-borne systems due to power and payload restrictions.

This paper describes an approach to sensor placement based upon an extension of techniques developed for sensor selection in monitoring as part of the SELMON project [Doyle et al. 91]. These techniques have two components, an information theoretic component and a model-based reasoning component. The information theoretic component captures knowledge about unusualness and informativeness of sensor data. The model-based reasoning component captures knowledge about how observed data indicates future potential system behavior. This paper focuses upon the model-based reasoning component of our sensor placement approach. In particular, we describe how model-based reasoning enables evaluation of four measures for evaluating potential sensor placements. *Sensitivity Analysis* suggests sensor placements which measure quantities which have the greatest impact upon the overall state of the system. *Cascading Alarm Analysis* suggests sensor placements which measure quantities whose changes have the potential to cause many alarms. *Potential Damage Analysis* suggests those sensor placements which measure quantities which are likely to cause permanent damage to devices in the system being monitored. *Teleological Analysis* suggests sensor placements which monitor quantities directly applicable to specified goal functionality of the system. Our approach uses a model-based simulation capability to evaluate how each sensor rates with respect to each of these measures over the behavioral space of the monitored system. These scores can then be used to generate a proposed sensor set.

This sensor placement evaluation capability provides a number of benefits. First, this evaluation capability will aid designers in the sensor placement task by facilitating evaluation of alternative sensor placements. In particular, this capability would provide a quantitative measure of tradeoffs in sensor placements which previously have been viewed only subjectively. A second benefit is that quantification of sensor placement measures will aid in design documentation by allowing quantitative justification for sensor placements. Third, the automated evaluation capability will facilitate assessment of the impact of system design changes upon sensor placements. Finally, as a fourth benefit, this sensor placement evaluation capability can be used to aid in sensor power planning. When the utility of a sensor depends greatly upon the operating mode of the monitored device, it may be possible to reduce overall sensor power consumption by powering certain sensor suites only in limited operating modes. Because our approach measures the utility of sensors in each system operating mode, it can assist in sensor power planning.

The remainder of this paper consists of three sections. The next section begins by describing how a model of the monitored system can be used to conduct a generalized simulation and how this simulation can be used to evaluate the four sensor placement criteria. The four sensor placement criteria are then described in greater detail. The following section describes the testbed domain - the Environmental Control and Life Support System (ECLSS) for Space Station Freedom (SSF). This section contains an example illustrating how the sensor placement criteria apply to the mutifiltration subsystem of SSF life support. The final section of the paper discusses outstanding issues regarding our approach to sensor placement, compares our approach to related work, and summarizes the key aspects of our approach.

Approach

Our approach to sensor placement is shown in Figure 1 and can be described generally as follows:

1. Given nominal behavioral models of the system and a causal simulation capability, generate a behavior space for the system.

2. Apply the sensitivity, cascading alarms, potential damage, and teleological analysis for system operation over these operating modes.

3. Compute sensor placement recommendations as those with highest scores from the analyses.

Our modelling uses a representation based upon [Doyle88]. In this representation, a model consists of a number of devices; each of which has a set of associated quantities. Mechanisms represent relationships between various quantities and are instances of physical processes. Simulation proceeds by executing mechanisms whose inputs change - possibly triggering other mechanisms. Mechanisms may change quantities, potentially with an associated delay. For example, fluid moving through a pipe may change the volume of fluid at the destination, but with a delay related to the size of the pipe and the flow rate.

We now describe the Sensitivity, Cascading Alarms, Potential Damage, and Teleological analyses. The Sensitivity, Cascading Alarms, and Potential Damage analyses use the model to simulate the effect of changes in quantity values. The effects of these changes are then analyzed to produce a Sensitivity, Cascading Alarms, or Potential Damage score. The Teleological Analysis uses a specification of the functional goal(s) for the system or subsystem along with an analysis of dependencies among mechanisms in the model to produce a sensor placement score.

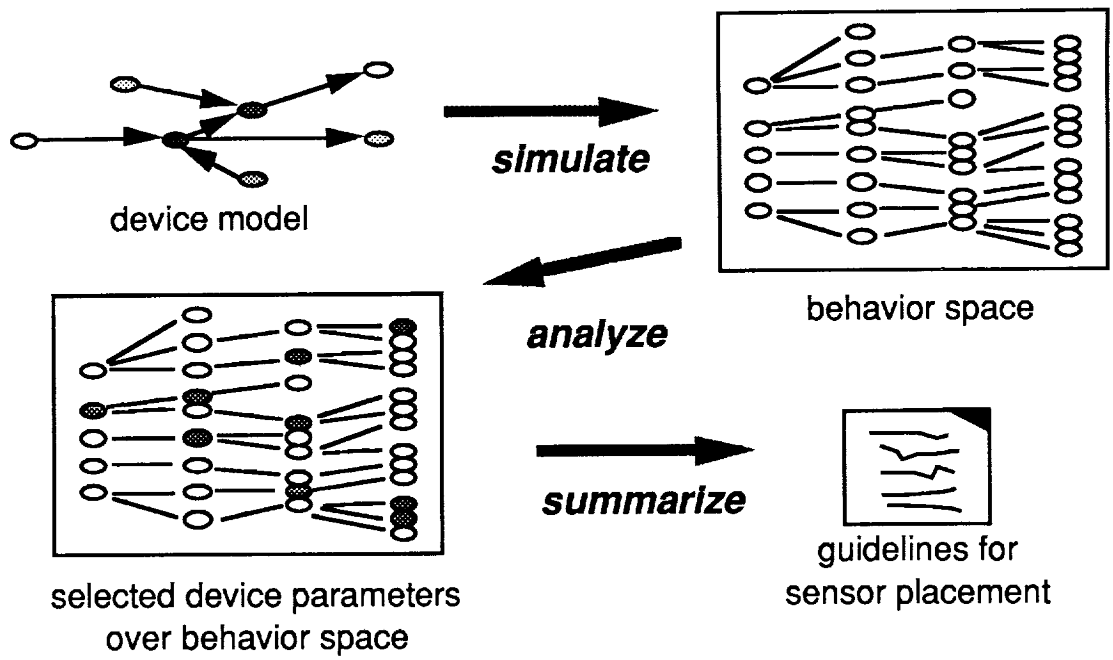

Figure 1: Model-based Simulation Approach to Sensor Placement

Sensitivity Analysis

Sensitivity Analysis measures the sensitivity of other quantities in the monitored system to changes in the current quantity. This measure depends upon information about "normal" magnitudes of change for the devices in question. For each normal operating mode of the system, the following procedure is performed. For each quantity Q $\in$ MonitorableQuantities (the set of all monitorable quantities in the model), determine nominal operating values and alarm ranges. Next compute a normalized change increase $\Delta Q+$ and decrease $\Delta Q-$ as the average amount of change between updates for that operating mode. Next, for each quantity Q, beginning with a simulation with all devices/sensors at nominal operating values, using simulation provided by the model, simulate a change ΔQ in Q, propagating this change to other quantities in AllQuantities (the set of all quantities in the model) as dictated by the model. For each such changed quantity Q' $\in$ AllQuantities, for each time the quantity changes during the simulation, collect a sensitivity score proportional to the amount of change in Q' from its normal value $Q'_{nominal}$ relative to alarm thresholds but also modified by a decreasing function of time[1]. This calculation captures the characteristic that delayed and less direct effects are more likely to be controllable and less likely to occur. Thus, a change which affected a quantity Q' but occurred slowly is considered less important. This simulation proceeds for a preset amount of simulated time. Then, for each changed quantity Q', take the maximum of the collected change score for that quantity. The sensitivity score for Q is the sum of these maximums for all the Q's. Thus, for each quantity Q, a simulated change produces a set of changescores for each other quantities in the model. The

[1]This can be viewed as an average $\partial Q'/\partial Q$ modified by a decreasing function of time elapsed and normalized for the alarm threshold for Q'.

sensitivity score for Q is the sum of the respective maximums of each of these sets[2]. The computation of the sensitivity scores is shown below:

Simulate a change $\Delta Q+$ or $\Delta Q-$ to Q beginning at time 0 and continuing to time ΔT (a user-supplied default).

For each change to a quantity Q' occurring at time T_{change}, compute a change score as follows.

let Q'_{new} be the new value for Q'

$$changescore(Q') = \frac{|Q'_{new} - Q'_{nominal}|}{|Q'_{alarm} - Q'_{nominal}|} \times \frac{(\Delta T - T_{change})}{\Delta T}$$

add this changescore to the set of collected changescores for Q'

let MaxChangeScore(Q') = the maximum of the set of collected changescores for Q'

$$\text{let } sensitivity(Q) = \sum_{Q' \in AllQuantities} MaxChangeScore(Q')$$

The overall sensitivity score for Q is then computed by summing the sensitivity scores for $\Delta Q+$ and $\Delta Q-$ weighted by relative frequency of increase vs. decrease for Q.

Cascading Alarms Analysis

Cascading alarms analysis measures the potential for change in a single quantity to cause a large number of alarm states to occur, thus causing information overload and confusion for operators. As with sensitivity analysis, cascading alarms analysis is performed for each operating mode of the monitored system. For a standardized amount of increase and decrease for each monitorable quantity Q, the effects of such a change are propagated throughout the system and the number of triggered alarms is counted. This standardized amount of change is different from the measure used in the sensitivity analysis as normal changes are not likely to produce cascading alarm patterns. The alarm count is then normalized for the total number of possible alarms. The weight of each alarm state triggered is also decreased as a function of the time delay from the initial change event to the alarm. This has the effect of focussing this measure on quickly developing cascading alarm sequences which are the most difficult to interpret and diagnose. The computation of cascading alarms scores is shown below.

Simulate a change $\Delta Q+$ or $\Delta Q-$ to Q beginning at time 0 and continuing to time ΔT (a user-supplied default); where $\Delta Q+$ and $\Delta Q-$ are functions of the distance between the nominal value for Q and the alarm value for Q in the increasing and decreasing directions respectively

[2]Quantities which do not change when Q is changed produce an empty set of changescores. We define the maximum of this empty set as 0 for the purpose of the sensitivity summation.

$$\text{let CascadingAlarm}(Q) = \frac{\sum_{Q' \in \text{ all quantities}} \text{InAlarm}(Q')}{\text{number of quantities } Q'}$$

where $\text{InAlarm}(Q') = (\Delta T - T_{alarm})/\Delta T$

if Q' entered an alarm range during the simulation and T_{alarm} is the earliest time Q' was in an alarm range and

$\text{InAlarm}(Q') = 0$

if Q' did not enter an alarm range during the simulation.

Potential Damage Analysis

Another measure is potential damage, which is measured in two parts - predictive potential damage and potential damage detection. Predictive potential damage measures the capability of the sensor to predict damage to devices in the system. For each device and quantity associated with that device, there is an associated operating range which is judged to be harmful to the device. Predictive potential damage analysis is performed by simulating a change in each monitorable quantity Q and scoring upon the basis of how many devices will enter harmful ranges due to the change in Q. Predictive potential damage analysis scores are moderated by the number of control points which may interdict the damage (defined as a mechanism directly influenced by a directly controllable quantity). For the causal path leading to the damaged device, for each mechanism which depends upon a control parameter, the potential damage score is reduced. The potential damage measure depends more critically upon domain-specific information beyond the schematic, as many of the potential damage scenarios involve device or subsystem interactions. The computation of potential damage scores is shown below.

Simulate a change $\Delta Q+$ or $\Delta Q-$ to Q beginning at time 0 and continuing to time ΔT (a user-supplied default).

$$\text{let PotentialDamage1}(Q) = \sum_{Q' \in \text{ all quantities}} \text{Damaged?}(Q')$$

where

$$\text{Damaged?}(Q') = \frac{(\Delta T - T_{alarm})}{\Delta T \times (\text{control} + 1)}$$

if Q' entered a damaging range during the simulation where T_{alarm} is the earliest time Q' was in a damage range and control is the number of control points in the causal chain leading to the damaging quantity value and

$\text{Damaged?}(Q') = 0$

if Q' did not enter a damage range during the simulation.

The second part of potential damage analysis is damage detection. In this measure, the model is used to simulate devices in the system entering damaging operating modes, and potential sensors are scored upon the basis of how much they change (in the same manner as the sensitivity analysis). Damage detection analysis is performed by propagating a change resulting in a device entering a damaging range, and measuring the resulting change in sensor reading as in the sensitivity analysis. Those sensors which change more significantly to indicate the damaging device state are scored higher by the damage detection analysis.

Let $\Delta Q'+$ or $\Delta Q'-$ be changes sufficient to cause Q' to enter a device damaging range.

Simulate a change $\Delta Q'+$ or $\Delta Q'-$ to Q' beginning at time 0 and continuing to time
ΔT (a user-supplied default).

$$\text{let PotentialDamage2}(Q) = \sum_{Q' \in \text{ all quantities}} \text{Changescore}(Q)$$

The final measure is teleological analysis, which does not use the model-based simulation capability. Instead, the teleological analysis examines the mechanism dependencies to produce a sensor placement score. In some cases, some aspects of the purpose of the monitored system can be specified in terms of quantities in the model, some of which may be monitorable quantities. The teleological analysis measure suggests measurements of quantities most closely linked to the functional goals of the system being monitored. In this measure, those quantities directly mentioned in the functional specification of the system are scored highest, those quantities directly influencing these quantities are scored next highest, etc. The exact computation of the teleological measure involves backtracing the causal graph. Directly monitorable quantities appearing in the goal description receive a score of 1. For each mechanism affecting the goal quantity, a teleology score inversely proportional to the number of such mechanisms is then divided equally among the inputs to the mechanism. Thus, if there are M mechanisms affecting a goal quantity, and one of these mechanisms has N inputs, each such input will receive a score 1/MN. Note that multiple causal influence paths will combine additively. While this process proceeds recursively for the mechanisms potentially influencing the inputs to this mechanism, each level is multiplied by 1/D where D is the number of mechanisms distant from the goal quantity.

The sensor placement scores computed by evaluating these four measures can be used in two ways. First, design engineers can use these scores directly to aid in their decision making process. Second, these scores can be combined in a utility function to rate the desirability of each potential sensor placement. The sensor placement task can then be viewed as maximizing the utility function within certain cost constraints (e.g. weight, power consumption, computing resources, cost, etc.). While we are currently investigating the sensor placement task within the context of the first approach (e.g. development of a sensor placement evaluation tool), we expect the resulting evaluations to be applicable to the second task.

Domain and Status

Our sensor placement approach is being tested upon the water reclamation subsystem of the Environmental Control and Life Support System (ECLSS) for Space Station Freedom. A model describing the behavior of the multifiltration (MF) subsystem in terms of fluid flow and heat transfer has been constructed. This model was developed via a combination of study of design documentation (i.e. schematics, etc.) and consultation with domain experts (e.g. the operators of the testbed). This model has been validated by comparison against actual data from the subsystem testbed undergoing evaluation at the Marshall Space Flight Center in Huntsville, Alabama. We are also in the process of extending our model to cover more of the ECLSS subsystems.

Figure 2 below shows the ECLSS multifiltration (MF) subsystem. The MF subsystem consists of two main parts - the sterilization loop and the unibed assembly. In the MF subsystem, the water first passes through a pump at the inlet to the system. Next, the water passes through a coarse filter before entering the sterilization loop. In the sterilization loop the water is heated in the regenerative heat exchanger and then by the in-line heater after point 3. The in-line heater has only a coarse temperature control and thus the water temperature at point 4 may differ as much as 10° F from the goal of 250° F. Within the sterilizer reservoir, the temperature of the water is maintained more accurately at 250°F for about 9 minutes. In the second portion of the subsystem, the water passes through a set of unibed filters designed to remove particulate contaminants from the water. Possible sensor types are flow rate, water pressure, temperature. Possible sensor locations are indicated in Figure 2 by numbered ovals.

Specified functional goals are:

1. maintain processed water at 250°F in sterilizer reservoir for 9 minutes; and
2. maintain water flow through the unibed of at least 15 mL/minute.

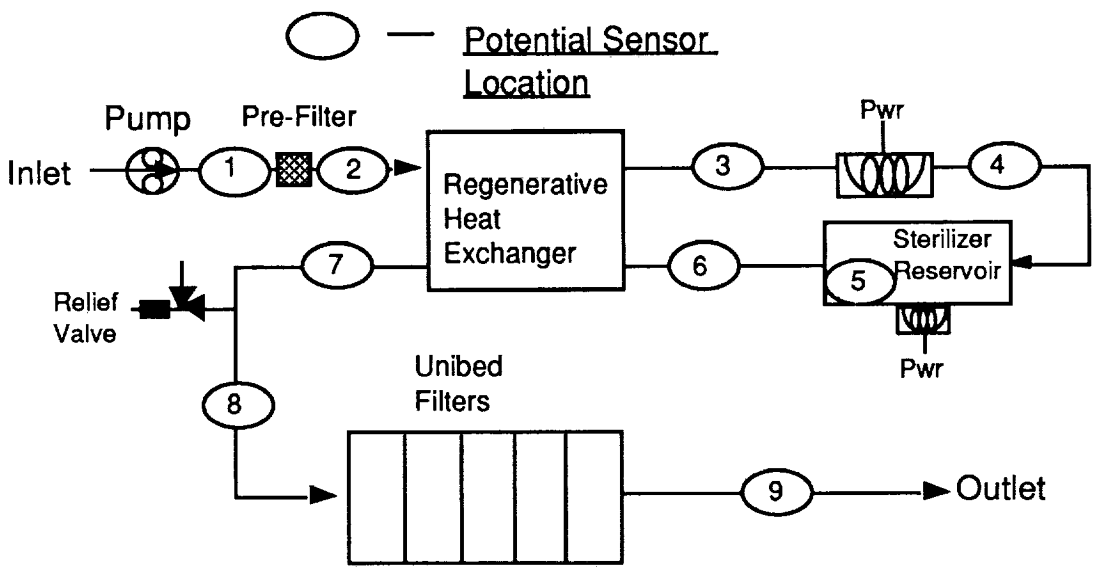

Figure 2: Multifiltration Subsystem

In this example, Damage Detection Analysis suggests placing a temperature sensor at point 4. This is because if the in-line heater overheats and enters a damaging temperature range, it would cause the water flowing through the in-line heater to be heated to a higher temperature than normal, thus causing a significant temperature increase at point 4.

Sensitivity Analysis scores highly the placement of a flow rate sensor anywhere in the MF because the flow rate affects many other quantities. The flow rate affects the temperature at each of the potential sensor placement points in that the flow rate affects the heat loss in the pipes. The flow rate also affects the speed of the temperature propagation from the fluid flow (e.g. delays for temperature changes).

Sensitivity Analysis also suggests the more specific placement of a pressure sensor near the relief valve at point 7. This is because the relief valve is pressure controlled; if the pressure at point 7 is above 40 psig, the relief valve will open and drastically change the system behavior. The opening of the relief valve would cause an immediate significant pressure loss, as well as significantly affecting flow in the MF subsystem.

Teleological Analysis suggests placing flow rate sensors to verify the flow of water through the unibeds as the flow rate is directly mentioned in the goal specification. Teleological Analysis also highly scores a flow rate sensor in the sterilizer reservoir, as this quantity determines the time spent by the water in the sterilizer reservoir. Teleological Analysis also scores slightly lower, flow rate sensors at any of the other locations, as in normal operation the flow rate is the same at all of the potential sensor locations. Finally, Teleological Analysis also suggests placement of a temperature sensor for the sterilizer reservoir, as this quantity appears in the functional goal specification of the system.

While the MF subsystem model has been completed and verified against actual testbed data, the sensor placement scoring algorithms are currently being implemented. When finished, the sensitivity, cascading alarms, potential damage, and teleological measures will then be tested on the current testbed configurations with domain experts evaluating the accuracy and utility of the sensor preference criteria.

Discussion and Conclusion

The research described in this paper is preliminary, hence there are a number of outstanding issues left to future work. One issue is that of the computational expense of simulation. While we are currently implementing a generalized simulation capability for all behavioral modes, this approach may be intractable for more complex systems Hierarchical simulation schemes and/or Monte Carlo sampling of the behavior space are currently being examined as approaches to dealing with this problem.

Another difficulty is determining how far forward to simulate for the sensitivity, cascading alarms, and potential damage analyses. Additionally, how large of a quantity change to propagate is another issue. Currently, both of these are parameters which may be changed by the user. However, ideally, the system would be able to determine appropriate values for these parameters.

Our work in sensor placement for monitorability is also relevant to issues in design for diagnosability. Because methods for ensuring diagnosability depend upon completeness of fault models, if fault models are incomplete, sensor placement based solely upon diagnosability criteria may make it difficult to detect and diagnose unforseen misbehaviors. Thus an approach to sensor placement based only on criteria of diagnosability may result in sensor configurations which do not support safe system operation. This is an important point about the difference between monitoring and diagnosis, or the difference between anomaly detection and troubleshooting. The goal of sensor placement for diagnosability is to ensure that sensor data to support the diagnosis of known faults will be available to operators. However, as the history of the Voyager mission tells [Laeser et al. 86], the potential for unforeseen faults with obscure manifestations always exists. There will be no substitute for years of experience and troubleshooting expertise on the part of domain experts in handling successfully these most difficult and potentially fatal cases.

In accordance with this observation, we are developing a comprehensive approach to sensor placement based upon both diagnosability and monitorability. This paper has described our work in design for monitorability in which the goal is to provide operators with high information content sensor data and/or sensor data which report on critical causal pathways in the device, without reference to fault models. Information-theoretic monitorability criteria capture knowledge about unusualness and informativeness of sensor data and the four model-based reasoning monitorability criteria described in this paper capture knowledge about how observed data indicates future potential system behavior. Although sensor placement based on monitorability does not provide a comprehensive solution to the problem of unforeseen faults, the intent is to provide experienced operators with the most generally informative sensor data and to avoid having sensor placements be wired into an inevitably incomplete set of fault models.

In a parallel effort, we are also examining the use of fault models to evaluate sensor placements from a diagnosability standpoint [Chien & Doyle91]. These methods complement the sensor placement for monitorability approach described in this paper. Thus, in those cases where fault models exist, a diagnosability analysis will incorporate such knowledge in sensor placement recommendations. For example, the types of interactions addressed by the Cascading Alarms and Potential Damage Analyses can be focussed by fault model information (by indicating which such interactions are likely to occur). Thus, fault model information is used in the diagnosability analysis. However, because monitorability analysis does not use fault models, it offers better coverage of unforseen faults.

Other work on sensor placement includes [Scarl91a, Scarl91b]. Scarl's work focuses upon two issues: 1) discrimination of faulty sensors (and hence faulty sensor data) from the occurrence of regular system faults; and 2) deriving minimal sensor sets to cover known fault sets. In contrast, our work focuses upon quantifying how well proposed sensors meet general monitorability criteria. Other relate work includes work on teleology and model-based reasoning [Sticklen et al. 89, Franke89, Sun & Sticklen90]. Sensor placement for monitorability is also related to design for testability [Wu88, Shirley88]. In this work, design constraints for digital circuits are derived from testability constraints. This work differs from our work in several respects. First, we are concerned with monitoring systems with continuous quantities. Second, issues of time criticality arise in our domain. And third, we are also concerned with monitoring in non-faulted modes.

This paper has described a model-based reasoning approach to evaluating sensor placements from the standpoint of monitorability. In this approach, potential sensor placements are evaluated using four criteria. Sensitivity Analysis suggests monitoring of quantities which, if changed, significantly affect overall system state. Cascading Alarms suggests monitoring of quantities which may lead to rapidly developing alarm sequences. Potential Damage suggests sensors which monitor quantities which predict or detect damage to devices in the monitored system. And

Teleological Analysis suggests monitoring of quantities more directly causally linked to the specified goal functionality of the system.

Acknowledgements

The research described in this paper was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. We would like to thank Jay Wyatt of Marshall Space Flight Center for innumerable discussions regarding the operation of the hygiene subsystem.

References

[Chien & Doyle91] S. A. Chien and R. J. Doyle, "A Model-based Reasoning Approach to Sensor Placement for Diagnosability," Internal Working Paper, Artificial Intelligence Group, Advanced Information Systems Section, Jet Propulsion Laboratory, California Institute of Technology, 1991.

[Doyle88] R.J. Doyle, "Hypothesizing Device Mechanisms: Opening Up the Black Box," Technical Report 1047, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, June 1988.

[Doyle et al. 91] R. J. Doyle, U. M. Fayyad, D. Berleant, L. K. Charest, L. S. Homem de Mello, H.J. Porta, and M.D. Wiesmeyer," Sensor Selection in Complex Systems Monitoring Using Information Quantification and Causal Reasoning" in Recent Advances in Qualitative Physics, B. Faltings and P. Struss (eds.), MIT Press, 1991.

[Franke89] D. Franke, "Representing and Acquiring Teleological Descriptions," Proceedings of the 1989 Workshop on Model Based Reasoning, August 1989, pp. 62-67.

[Laeser et al. 86] R. P. Laeser, W. I. McLaughlin, and D. M. Wolff, "Engineering Voyager 2's Encounter with Uranus", *Scientific American*, 255(5):36-45, November 1986.

[Scarl91a] E. Scarl, "Diagnosability and Sensor Reduction," Proceedings of the Workshop on AI, Simulation, and Planning in High Autonomy Systems, Cocoa Beach, FL, 1991.

[Scarl91b] E. Scarl, "Monitoring and Diagnosis: Stress from Weakened Environmental Knowledge," Proceedings of the 1991 IEEE Conference on Robotics and Automation, Sacramento, CA, 1991.

[Sticklen et al. 89] J. Sticklen, B Chandrasekaran, and W. Bond, "Applying a Functional Reasoning Approach for Model-based Reasoning," Proceedings of the 1989 Workshop on Model Based Reasoning, August 1989, pp. 165-176.

[Sun & Sticklen90] J. Sun and J. Sticklen, "Steps Toward Tractable Envisionment via a Functional Approach," Proceedings of the 1990 Workshop on Model-based Reasoning, August 1990, pp. 50-55.

[Shirley88] M. Shirley, "Generating Circuit Tests by Exploiting Designed Behavior," Technical Report 1099, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, December 1988.

[Wu88] P. Wu, "Test Generation Guided Design for Testability," Technical Report 1051, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA, May 1988.