AN OPERATOR MODEL-BASED FILTERING SCHEME*

R. S. Sawhney

University of Tennessee
Knoxville, Tennessee 37996

J. C. Schryver

Oak Ridge National Laboratory
Oak Ridge, Tennessee 37831

H. L. Dodds

University of Tennessee
Knoxville, Tennessee 37996

To Be Presented At:

ANS Topical Meeting on "Advances in Human Factors Research on Man-Computer Interactions: Nuclear and Beyond"

Opryland Hotel
Nashville, Tennessee

June 10-14, 1990

*Research sponsored by the Office of Technology Support Programs, U.S. Department of Energy under Contract No. DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

AN OPERATOR MODEL-BASED FILTERING SCHEME

R. S. Sawhney
Dept. of Industrial Engineering
University of Tennessee
Knoxville, Tennessee
[615] 974-3333

J. C. Schryver
Oak Ridge National Laboratory
Post Office Box 2008
Oak Ridge, Tennessee
[615] 574-4710

H. L. Dodds
Dept. of Nuclear Engineering
University of Tennessee
Knoxville, Tennessee
[615] 974-2525

ABSTRACT

This paper presents a diagnostic model developed at Oak Ridge National Laboratory (ORNL) for off-normal nuclear power plant events. The diagnostic model is intended to serve as an embedded module of a cognitive model of the human operator, one application of which could be to assist control room operators in correctly responding to off-normal events by providing a rapid and accurate assessment of alarm patterns and parameter trends. The sequential filter model is comprised of two distinct subsystems - an alarm analysis followed by an analysis of interpreted plant signals. During the alarm analysis phase, the alarm pattern is evaluated to generate hypotheses of possible initiating events in order of likelihood of occurrence. Each hypothesis is further evaluated through analysis of the current trends of state variables in order to validate/reject (in the form of increased/decreased certainty factor) the given hypothesis.

1. INTRODUCTION

In a highly complex person-machine system such as a nuclear power plant, the operator's capabilities become a critical issue in maintaining the plant in a normal operating condition. The operator receives numerous stimuli and is expected to respond both expediently and in a correct manner to diffuse off-normal events. This expectation is based on the assumption that the operator is capable of efficiently performing the tasks illustrated in Fig. 1: (a) perceive information, (b) process information, and (c) perform the appropriate action in response to the processed information.

The stress resulting from rapid presentation of stimuli within a short period of time will test the human capabilities to perform effectively. For example, one limiting characteristic is the operator's ability to retrieve information from long-term memory. Deficiencies in knowledge retrieval can result in human error. A model that filters plant-based signals to aid the operator in the information processing and decision-making phase of Fig. 1 is the first step in developing a comprehensive decision support system.

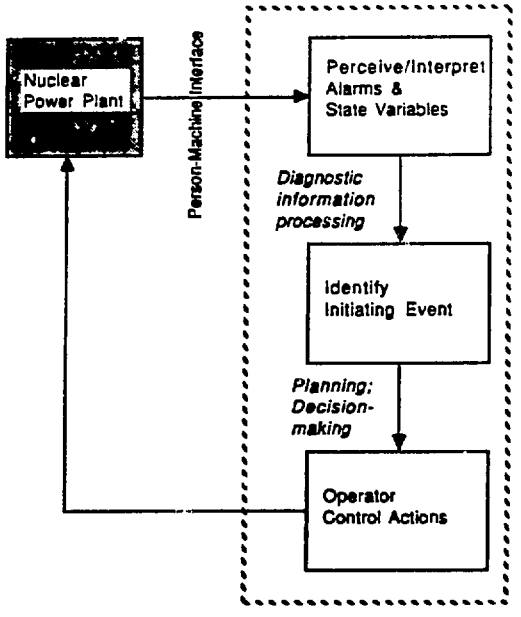

FIGURE 1. Person-machine system

2. BACKGROUND

This research was performed to support the development of the Integrated Reactor Operator/System (INTEROPS) model[1] under the auspices of the Advanced Controls Program at ORNL. INTEROPS is a model of cognitive functions and discrete control actions required to operate a commercial reactor module of the Power Reactor Inherently Safe Module-Advanced Liquid Metal Reactor (PRISM-ALMR)[2] design. The model is coupled with a PRISM-ALMR thermal-hydraulics plant simulation. The INTEROPS model includes two types of components: a SAINT simulation module (a task oriented network simulation model of the human operator)[3], and knowledge-based simulation modules which include a plant filtering model that generates and verifies initiating event hypotheses.

Many alarm filtering models have been described in the literature. Examples are DuPont's event tree-based diagnosis of multiple alarms[4] and ORNL's alarm filtering system which is based on emphasis and deemphasis of alarms.[5] However, the rule-based diagnostic model - illustrated in Fig. 2 - is unique because it decomposes the analysis into two subsystems: an alarm analysis model to generate the initiating event hypothesis and a parameter analysis model to verify the hypotheses. Critical initiating events are identified through an alarm-initiating event network. The ranking of initiating events is determined by the certainty factor associated with each initiating event. The alarm analysis model modifies these certainty factors based on information from both annunciated (active) and unannunciated (latent) alarms while the parameter analysis modifies the certainty factors based on information from interpreted plant signals. Input data are also associated with certainty factors. For example, alarm annunciation causes the maximum certainty value to be affixed to the alarm label. Certainty factors associated with unannunciated alarms grow exponentially with time toward "negative certainty." Plant signals are interpreted by associating a verbal tag with a time series. For example, the tag "DIA" describes a time history which began by decreasing, then increasing, and finally asymptoting to a constant value. Plant parameter trend certainty factors decay exponentially between updates.

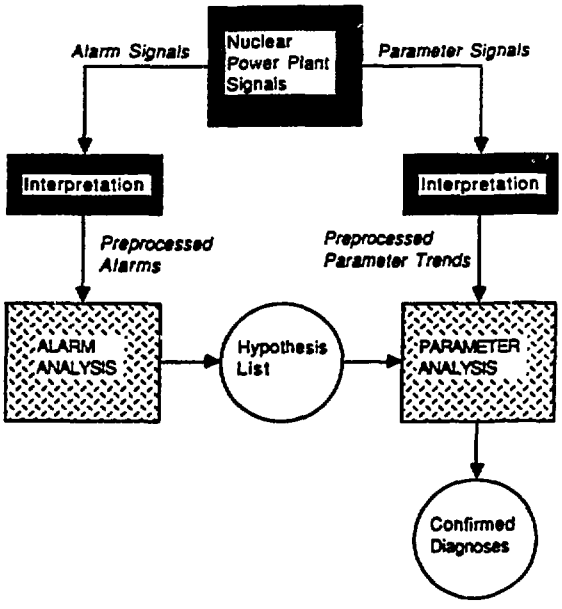

FIGURE 2. Overview of a rule-based diagnostic model

The method of confirmation and disconfirmation for inexact reasoning used in the medical expert system, MYCIN, was used to represent the effects of uncertain knowledge in advantages of ease of implementation, and, according to Lee, Grize, and Dehnad,[6] strong intuitive appeal. The latter characteristic supports confirmation/disconfirmation as a plausible model of human diagnostic reasoning. Instead of working with probability density functions, the underlying metric is the "certainty factor," which expresses the degree of confidence in a hypothesis.

The manner in which the model reasons with uncertain knowledge is a central issue in both the alarm and parameter analysis subsystems. The process of combining two certainty factors obtained from incrementally acquired evidence is used to modify an initiating event certainty factor in order to confirm or reject the initiating event as the plant transient source. The combining function for certainty factors (when both are positive) is given by the following equation:

$$CF\,[h,e_1\&e_2] = CF\,[h,e_1] + CF\,[h,e_2] \cdot (1-CF[h,e_1]) \qquad (1)$$

where h is the hypothesis or candidate initiating event, and e_1 and e_2 are the old and new evidence, respectively. The CF value ranges between -1 and +1. There are three other variations of this equation to account for other cases which arise depending on whether the old and new evidence is confirming or disconfirming.

3.1 Alarm Analysis

Alarm analysis is performed by a rule-based system which can be represented graphically as a tri-level network. The bottom layer shows the alarm input. The intermediate layer reveals the extent to which major functional subsystems are implicated in the fault sequence. The top layer is the output of the rule-based system and contains causes or initiating events. A reduced sample alarm-initiating event network is illustrated in Fig. 3.

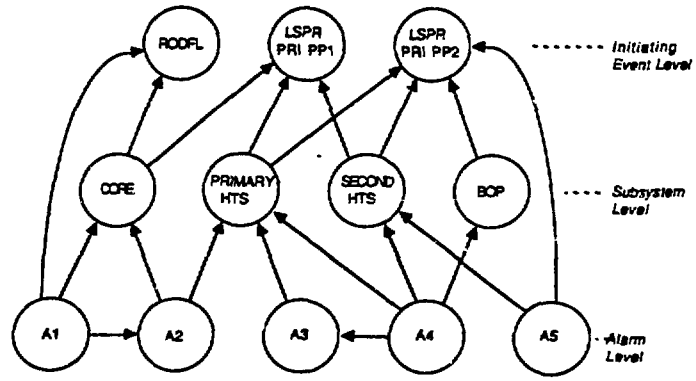

FIGURE 3. Hypothetical alarm-event rule-based network

The implication of this diagram is that alarm patterns are domain-specific; each alarm pattern implies a specific initiating event within the limits of certainty. For example, alarms A1 and A5, respectively, imply that the initiating event is rod failure and loss of power to primary pump 2. With less certainty both alarms also implicate the loss of power to

primary pump 1. Alarms A2 and A3 indicate that the problem is in the core and primary heat transport subsystems which lead to rod failure, and loss of power to primary pumps. The arrows in Figure 3 illustrate the following relation types which are the basis for "rules" in the alarm analysis:

1. The functional subsystem domain (core, primary heat transport system, secondary heat transport system, balance-of-plant) for a given alarm.

2. The initiating event domain (rod failure, loss of power to pump) for an alarm.

3. The initiating event domain for a functional subsystem.

4. The alarm-to-alarm mappings.

The rule-based model would function as efficiently without the intermediate subsystem level of processing. However, its implementation as an additional filter allows the model to use fewer rules with little loss of accuracy, and perhaps some gain in flexibility. The concept of hierarchical filters does seem to have some plausibility for human diagnostic reasoning.

The alarm analysis accomplishes its objectives by providing mappings, along with associated certainty factors, among the three functional levels. The alarm analysis modifies initiating event certainty factors in five major steps to identify hypotheses for parameter analysis:

1. Establish alarm input values.

2. Modify alarm certainty factors with the alarm-alarm mapping.

3. Calculate subsystem certainty factors with the alarm-subsystem mapping.

4. Establish initiating event certainty factors with the alarm-event mapping.

5. Modify initiating event certainty factors with the subsystem-event mapping.

The certainty factors of annunciated alarms remain at maximum value. The amount of information content in unannunciated alarms is initially zero and increases exponentially with time. That is, the model is fairly certain that a quiescent alarm will not suddenly annunciate if considerable time has elapsed since failure detection. This information is important for "boundary definition" in pattern recognition during alarm processing. The alarm-alarm mappings show the expectancies generated by the model as part-patterns are recognized, and the part-whole inferences are completed. The model may notice that the alarms are concentrated in one or two regions of the system while processing the alarm-subsystem mapping. Initiating event certainty factors are calculated by analyzing two classes of antecedents: alarm certainty factors and subsystem certainty factors.

The alarm analysis is unique in the sense that it generates hypotheses based on more highly processed information. It incorporates information not utilized by other models in diagnosing initiating events for off-normal scenarios. In particular, there are two types of information which hold the opportunity to improve the prioritization of hypotheses. First, the model has the ability to use alarm-alarm mapping to modify alarm confidence factors based on expectations derived from the annunciated alarm patterns. Second, the model performs an unannunciated alarm analysis to verify and modify the results of the annunciated alarm analysis.

The information in the alarm-alarm mapping was developed by comparing the responses of the plant model to a complete transient set with respect to each alarm pair. The output of this exercise was used to construct an alarm correlation matrix. Positive correlations indicate a tendency for an alarm to annunciate in the presence of another annunciation. Negative correlations indicate a reduced tendency for an alarm to annunciate, given another alarm has already annunciated. Only extreme-value correlations were represented in the alarm-alarm mapping in the rule-based network; nevertheless, they provided very rich data sources for the purposes of alarm analysis.

Alarm-alarm domain interactions modify alarm certainty factors based on "expectations" derived from the annunciated alarm patterns. For example, if there is a positive correlation between A4 and A6, and A4 has annunciated but A6 has not, confidence in the pending annunciation of A6 will be increased. A set of annunciated alarms is utilized to modify the certainty factor associated with each alarm according to the following steps:

1. For each member of the annunciated alarm list, create:
   (a) Pos-List - a list of all alarms that should be annunciated.
   (b) Neg-List - a list of all alarms that should not be annunciated.

2. For each match between the annunciated alarm list and Pos-List, the certainty factor for the alarm is increased.

3. For each match between the annunciated alarm list and Neg-List, the certainty factor for the alarm is decreased.

Information contained in unannunciated alarms may also be utilized to improve the accuracy of the hypotheses. The assumption is that there is valuable information content in unannunciated alarms that may be useful in generating more accurate hypotheses. Information from alarms that have not annunciated is used in the following manner. Hypotheses are generated from alarms that have not annunciated exactly in the manner in which hypothesis were developed based on annunciated alarms. If the hypothesis matches the hypothesis generated from annunciated alarms, the certainty factor of that hypothesis is increased. If the information is disparate, the certainty factor is reduced.

The output from the alarm analysis model (a list of initiating events given in descending order of the certainty factors) is utilized as input to the parameter analysis model which analyzes observed parameter trends to generate a new set of certainty factors for those initiating events selected as hypotheses by the alarm analysis subsystem.

3.2 Parameter Analysis

The objective of the parameter analysis is to confirm hypotheses using plant-based state variables (e.g., reactor thermal power, core outlet temperature). The critical input data for this analysis are a list of parameter trends. Note that only the trends of a given parameter are evaluated rather than numeric values, because the interpreted data contains the essential information which is processed in human diagnostic reasoning. The trend descriptions can be represented in a tree diagram as in Fig. 4.

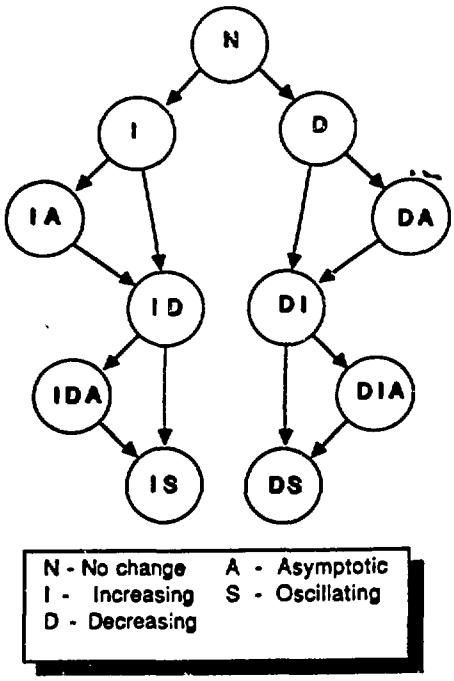

FIGURE 4. Parameter trend evolution tree

The parameter analysis model performs an evaluation of 31 plant parameters. The modification of certainty factors for the hypothesis list is conceptually simpler than the alarm analysis. The rule-based system can be represented by a simple two-layer network diagram. There is only one type of rule, the mapping between parameter trends and initiating events.

The process of confirmation employed by the parameter analysis is more like "backward chaining" than the "forward chaining" which characterizes the alarm analysis network. The latter analysis technique is data-driven - all alarm information is used. However, the use of information contained in state variables is selective. Only information which may be used to support an active hypothesis will be integrated into the chain of inference. The kind of narrowing of the data stream that results from this confirmation bias can be an efficient and powerful analytic technique. However, it is also a potential error source.

The following example is given to illustrate how information theory is invoked to determine the certainty factors associated with the mapping between parameter trends and candidate diagnoses. Suppose we want to estimate the certainty factor associated with the inference linking the hypothesis (h) of a recirculation pump trip and the evidence (e) that the turbine/bypass flow rate is equal to "DA." Referring to the trend evolution tree in Fig. 4, we notice that the fault set (F_o) as we enter the tree at top-level, or "N," contains all hypotheses. A series of discrimination operations are performed on F_o as a trend description evolves. The first discrimination produces "D," leaving the first-order fault set F_1. The first-order fault set contains all initiating events which initially cause turbine/bypass flow rate to decrease. The second discrimination at "DA" produces F_2. Notice that in general, it must be the case that $F_o \supseteq F_1 \supseteq \ldots \supseteq F_n$. Now let the correctness of the i^{th} discrimination be denoted by:

$$z_i = \begin{cases} -1 \text{ if } h \notin F_i \\ +1 \text{ if } h \in F_i \,. \end{cases}$$

In the present example all discriminations are correct, so $z_1 = 1$ and $z_2 = 1$. However, if the turbine/bypass flow rate began by increasing, then z_i would equal -1 for all i. Eq. (2) gives the method for calculating the certainty factor of h for the n^{th} discrimination:

$$CF[h,e \mid F_o,F_1,\ldots,F_n] = \sum_{j=1}^{n} z_j(\ln F_{j-1} - \ln F_j)/\ln F_o \qquad (2)$$

Eq. (2) guarantees that the contribution of each discrimination to the certainty factor is proportional to the amount of information contained in the discrimination. The correctness of the discrimination determines whether the certainty factor is incremented or decremented.

The output of the parameter analysis is a list of all hypotheses which are confirmed when the associated certainty factor exceeds a threshold value.

4. IMPLEMENTATION

The rule-based diagnostic model described in this paper is being implemented as a common LISP program on a VMS-based Vax superminicomputer. The main steps in the model are performed by function calls which reference the DEFUN macro. This macro allows any new function to be defined. Connecting a sequential series of DEFUN macros allows for the development of the rule-based system. The plant-based knowledge is structured utilizing the DEFSTRUCT macro. There are two primary advantages to this approach. First, there is simplicity of implementation. Second, the model has great flexibility with respect to addition or deletion of alarms, subsystems, parameters, initiating events, and rules.

The alarm analysis subsystem has been implemented and undergone some testing with PRISM-ALMR transients. The implementation of the parameter analysis has been defined and is currently being coded for testing.

5. CONCLUSION

Preliminary results of the alarm analysis based on plant model output from ten transients indicate that hypothesis generation performance is quite satisfactory.[7] The alarm analysis has not been tested under conditions of multiple failure, which would provide a stronger test of its pattern classification capability. The parameter analysis is expected to provide very good classification under single failure conditions prior to intervention of the control and protection systems, or compensatory control actions of the human operator. These compensatory events tend to disturb parameter time histories or trends, altering their trajectories relative to the trajectories expected through propagation of the effects of the original fault. Future model development efforts are expected to focus on theory expansion of the analysis of parameter trends.

REFERENCES

1. J. C. Schryver, "Operator Model-Based Design and Evaluation of Advanced Systems: Computational Models," *Proceedings of the IEEE Fourth Conference on Human Factors and Power Plants*, pp. 121-127, Monterey, CA. (June 1988).

2. Y. Dayal, "Advanced PRISM Plant Control System," *Proceedings of the 7th Power Plant Dynamics, Control and Testing Symposium*, pp. 1.01-1.13, Knoxville, TN. (May 1989).

3. G. P. Chubb, K. R. Laughery, Jr., and A. A. Pritsker, "Simulating manned systems," pp. 1298-1327 in G. Salvendy (Ed.), *Handbook of Human Factors*, New York: John Wiley & Sons, 1987.

4. K. L. Gimmy, "Plant Experience with an Expert System for Alarm Diagnosis," *Proceedings of the Society of Computer Simulation*, San Diego, CA. (January 1986).

5. J. T. Robinson, P. J. Otaduy, "An Expert System for Alarm Diagnosis and Filtering," *Proceedings of the ANS Topical Meeting on Artificial Intelligence and Other Computer Applications in the Nuclear Industry*, Snowbird, UT. (August 1987).

6. N. S. Lee, Y. L. Grize, and K. Dehnad, "Quantitative Models for Reasoning Under Uncertainty in Knowledge-Based Expert Systems," *International Journal of Intelligent Systems*, Vol. II, pp. 15-38 (1987).

7. R. S. Sawhney, J. C. Schryver, H. E. Knee, and H. L. Dodds, "A Hypothesis Generation Model of Initiating Events for Nuclear Power Plant Operators," *Proceedings of the 1989 ANS Winter Meeting*, San Francisco, CA. (November 1989).