Automatica, Vol. 19, No. 6, pp. 613–625, 1983
Printed in Great Britain.

0005-1098/83 $3.00 + 0.00
Pergamon Press Ltd.

Models of Human Problem Solving: Detection, Diagnosis, and Compensation for System Failures*

W. B. ROUSE†

A review of the human's role in detecting, diagnosing, and compensating for system failures permits the synthesis of a multilevel, rule-based, pattern recognition oriented model of human problem solving in such situations.

Key Words—Man-machine interaction; human problem solving; detection; diagnosis; compensation; system failures.

Abstract—The role of the human operator as a problem solver in man–machine systems such as vehicles, process plants, transportation networks, etc. is considered. Problem solving is discussed in terms of detection, diagnosis, and compensation. A wide variety of models of these phases of problem solving are reviewed and specifications for an overall model outlined.

INTRODUCTION

MAN–MACHINE interaction has been a topic of formal study for well over 50 years. The earliest investigations focused on the environment as it affected the human operator's safety and ability to perform his job. Later investigations began to consider also the design of equipment in terms of identifying possible limitations and devising potential enhancements of operator performance. Much progress has been made in these areas, although a great deal remains to be accomplished.

More recently, the impact of automation has come to be of increasing importance. In aircraft, ships, process plants, transportation networks, and other large-scale systems, more and more control loops that were once closed manually are now automatically controlled. As a result the human operator is becoming more of a monitor and supervisor of automation (Sheridan and Johannsen, 1976).

The possibility of failures is the primary reason for having human monitoring of automatically controlled processes. If hardware and software failures could not occur and if the automation were capable of handling all contingencies, then human operators would be unnecessary. However, failures and design limitations are quite possible and therefore, a primary task of the human operator is to detect these events and deal with them appropriately. If current trends continue, this task will come to dominate the human's responsibilities (Rasmussen and Rouse, 1981).

It seems reasonable to make the general claim that the manual activities of the human operator will increasingly be supplanted by problem-solving activities. The objective of this paper is to review models of human problem solving with emphasis on models useful for describing and predicting human behavior and performance for the purposes of design and evaluation. While a brief review of the general area of human problem solving is presented, the models considered in most detail are those which are directly applicable to situations involving man–machine interaction in detecting, diagnosing, and compensating for system failures. The combined results of this brief general overview and the detailed review of the most relevant models are used as a basis for outlining specifications for an overall model of human problem solving.

GENERAL BACKGROUND

Much of the literature in the general area of human problem solving emphasizes the pattern recognition nature of human behavior in problem solving tasks. This position is argued from both a physiological basis (Albus, 1970) and using notions such as cognitive economy (Hormann, 1971). The pattern recognition need not be a concise one-to-one mapping. Familiar scripts (Schank and Abelson, 1977) or frames (Minsky, 1975) may evoke a sense of having seen a particular type of problem before. Of course, particular instances can also be recalled (Neimark and Santa, 1975).

The use of pattern recognition or visually oriented approaches to problem solving has been advocated for a variety of domains including chess (Chase and Simon, 1973; Nievergelt, 1977), elec-

* Received 13 April 1982; revised 19 July 1982. The original version of this paper was presented at the IFAC/IFIP/IFORS/IEA Conference on Analysis, Design, and Evaluation of Man–Machine Systems which was held in Baden-Baden F.R.G. during September 1982. The published proceedings of this IFAC meeting may be ordered from Pergamon Press Ltd, Headington Hill Hall, Oxford OX3 0BW, U.K. This paper was recommended for publication in revised form by editor A. Sage.

† Georgia Institute of Technology, School of Industrial and Systems Engineering, Center for Man-Machine Systems Research, Atlanta, GA 30332, U.S.A.

tronic troubleshooting (Dale, 1958; Burroughs, 1979), mental arithmetic (Hayes, 1973), and analogical problem solving (Sternberg, 1977). This view of analogies is rather interesting because it involves a double mapping; one to recognize the analogy and one to transform the solution.

Modes of problem solving

Not all problems can be solved by a direct mapping from observations to solution. Thus, modes of problem solving other than pattern recognition may be required. In an effort to describe possible multiple modes of problem solving, various dichotomies have been suggested. Examples include pattern recognition vs heuristics (Gerwin, 1974; Burroughs, 1979), intuition vs analysis (Peters, Hammond and Summers, 1974; Simonton, 1975), remembering vs solving (Rumelhart and Abrahamson, 1973; Jacoby, 1978), retrieval vs search (Atwood and co-workers, 1978), imagistic vs linguistic strategies (Wood, Shotter and Godden, 1974; Sternberg and Weil, 1980), and symptomatic vs topographic strategies (Rasmussen and Jensen, 1974; Rasmussen, 1978, 1981).

The choice of mode of problem solving can be highly influenced by the way in which the problem is represented. Perceptual cues, even if they are irrelevant, can lead to a pattern recognition mode of behavior (Dale, 1958; Peters, Hammond and Summers, 1974). On the other hand, representations that preclude or inhibit pattern recognition may lead to a more analytical or heuristic approach (Peters, Hammond and Simmons, 1974; Gerwin and Newstad, 1977) and, at least in the form of flow charts or functional diagrams, such representations have been shown to improve some aspects of problem solving performance (Mayer, 1975; Brooke and Duncan, 1980). Some seemingly helpful forms of representation such as color coding (Neubauer and Rouse, 1979) and special formats (Brooke and Duncan, 1981) may, however, have surprisingly little effect, especially for highly practiced problem solvers. Another effect of experience is that humans tend to fixate on one form of representation even when multiple forms are available (Polich and Swartz, 1974).

Nature of expertise

Since multiple modes of problem solving are possible, it is quite natural to wonder if some modes are better than others. One way to approach this issue is to compare behavior of novices and experts. Many authors have argued that expertise is synonymous with highly developed pattern recognition abilities (Chase and Simon, 1973; Atwood and co-workers, 1978; Dreyfus and Dreyfus, 1979). Others have presented results that indicate expertise to be related to particular strategies (Goldbeck and co-workers, 1957; Wood, Shotter and Godden, 1974; Simon and Reed, 1976; Hayes-Roth and Hayes-Roth, 1979).

Depending on one's definition of pattern recognition, these two perspectives of expertise may or may not be conflicting. If one views expertise as being gained solely by the acquisition of a large repertoire of context-specific patterns, then strategy may be an irrelevant concept (Dreyfus and Dreyfus, 1979). On the other hand, if expertise in pattern recognition includes an ability to recognize useful context-free structural patterns in problems, then changes in strategy that reflect expertise can be described as changes in the perceived usefulness of structural patterns (Rouse, Rouse and Pellegrino, 1980).

The distinction between patterns of context-specific observations and patterns of context-free structures is central to the model outlined later in this paper. One school of thought emphasizes the dominance of context (Newell and Simon, 1972; Chase and Simon, 1973; Bree, 1975) while others give more credence to context-free aspects of problem solving (Kearsley, 1975; Mason, Bramble and Mast, 1975; Brooke, Duncan and Cooper, 1980) and the importance of structure (Loftus and Suppes, 1972; Malin, 1979; Rouse, 1981b). This issue can be clarified and partially resolved by reviewing the results of a variety of transfer of training studies.

Transfer of training

If context dominates problem solving, then transfer of training should be negligible between problems that are structurally similar but contextually different. While there is some evidence of such a lack of transfer (Smith, 1973), most results indicate positive transfer of training (Shepard and co-workers, 1977; Siegler, 1977; Luger and Bauer, 1978; Rouse, 1981b). Thus, structure is clearly an important aspect of problem solving. However, context is also important and does probably dominate when the human is in a familiar problem solving environment.

Considering the general area of training, it is interesting to contrast the modes of problem solving that various training methods promote. There are methods that emphasize context-specific pattern recognition (Duncan and Shepard, 1975; Towne, 1981; Johnson, 1981), methods that stress inferential (i.e. searching) strategies with respect to particular systems (Glaser, Darmin and Gardner, 1954; Landa, 1972; Brown, Burton and Bell, 1975; Freedy and Lucaccini, 1981; Hunt and Rouse, 1981), and methods that promote context-free search strategies (Rouse, 1981b). Recently, it has been argued that a mixture of methods is probably the best overall approach to training (Johnson and Rouse, 1982a, b; Rouse, 1982).

Models

Many of the above results have motivated the development of models of human problem-solving behavior. Some of the earlier efforts in this area compared human performance to optimal half-split strategies (Goldbeck and co-workers, 1957; Dale, 1958; Mills, 1971) or to time-optimal strategies (Stolurow and co-workers, 1955). There were also efforts to model the problem solver as a Bayesian information processor (Bond and Rigney, 1966; Kozielecki, 1972). These modeling endeavors have typically shown that the human is only optimal for simple problems unless extensive training is provided.

More recently, emphasis has come to be placed on the process (i.e. strategy) rather than the product (i.e. results) of problem solving (Gregg and Simon, 1967). In other words, concern has shifted to modeling human behavior (i.e. what the human does) rather than human performance (i.e. how well the human does). A variety of methodologies useful for developing process models have emerged (Newell and Simon, 1972; Waterman and Hayes-Roth, 1978; Rouse, 1980). As a result of this trend, considerable attention has been devoted to the concept of strategy (Wason and Johnson-Laird, 1972; Simon, 1975; Simon and Reed, 1976; Johnson, 1978). Some effort has been devoted to developing performance measures that are sensitive to differences in strategy (Duncan and Gray, 1975; Brooke and Duncan, 1980; Hunt and Rouse, 1981; Henneman and Rouse, 1982).

Process models have emerged for a wide range of tasks including chess (Newell and Simon, 1972), fitting of mathematical functions (Huesman and Cheng, 1973), water jug problems (Atwood and Polson, 1976), missionaries and cannibals (Jeffries and co-workers, 1977), errand planning (Hayes-Roth and Hayes-Roth, 1979), fault diagnosis (Rouse and co-workers, 1980; Rouse and Hunt, 1981; Hunt and Rouse, 1982), and many others. A common feature of these models is the extensive use of rule-based strategies (as opposed to algorithmic optimization) with, in a few cases, a mixture of probabilistic or fuzzy choices or transitions among modes. Another important aspect of these models is that they can actually perform the task of interest; this cannot be said of many of the earlier product-oriented models of human performance.

Summary

This brief review of the general area of human problem solving has served to point out several concepts and issues that are important to modeling human detection, diagnosis, and compensation for system failures. Perhaps the most important concept is the dichotomy between context-specific pattern recognition and structure-oriented searching, and the relationship of this dichotomy to the nature of expertise. It appears that training methods, forms of problem representation, and use of aids are important determinants of the modes of problem solving chosen by humans, at least initially.

While many of the models of human problem solving reviewed in this section are not directly relevant to the type of problem solving of interest in this paper, the notion of modeling the process rather than the product of problem solving is very important. Further, as later discussions will illustrate, much of the rule-based modeling methodology underlying these models has been quite useful for developing models that focus on detection, diagnosis, and/or compensation. Several of these models will now be reviewed.

MODELS OF DETECTION

Detection is defined as the process whereby the human operator decides that an event has occurred. There are four types of model of human performance in event detection. One type is based on signal detection theory; another type utilizes thresholds for error and error rate; another employs the residuals of a Kalman filter within a sequential decision theory algorithm; a final type is based on pattern recognition methods. These types of model are summarized in Table 1.* The following discussion elaborates on the summaries in this table.

Signal detection theory

Signal detection theory (see entry 1, Table 1) has been used extensively to describe the results of experimental studies of the human's abilities to detect infrequent signals in the presence of noise (Sheridan and Ferrell, 1974). The theory assumes that the human forms a likelihood ratio in terms of the conditional probability of observed data given there is a signal divided by the conditional probability of the observed data given there is only noise. This likelihood ratio is then compared to a threshold which is a function of *a priori* probabilities, values of correct responses and costs of incorrect responses. If the likelihood ratio exceeds this threshold, the theory predicts that the human will report the detection of a signal.

Results of signal detection studies are typically expressed in terms of 'hits' and 'false alarms'. The probability of a hit is plotted against the probability of false alarm. The resulting plot is called a relative

* The value of N shown in Tables 1 and 2 denotes the number of experimental subjects with which the model was compared.

TABLE 1. MODELS OF HUMAN FAILURE DETECTION

Model \ Attributes	Basic Approach	Key Assumptions	Types of Failure	Experimental Results
1. Signal Detection Theory	detection threshold on likelihood ratio	separability of detectability and decision criterion	presence of a single known event obscurred by noise; monitoring	typically used to describe hit vs. false alarm rates
2. Miller and Elkind (1967)	detection threshold on variance of changes in error rate	known set of possible failures	changes in gain and polarity for 1^{st} order system; compensatory tracking	reasonable predictions of detection time and some false alarms; N=3
3. Phatak and Bekey (1969)	detection thresholds on error and error rate	known set of possible failures	changes between 2^{nd} and 4^{th} order aircraft dynamics; compensatory tracking	reasonable qualitative comparisons of detection time and thresholds; N=1
4. Niemela and Krendel (1975)	detection thresholds on error and error rate	known set of possible failures	changes in polarity for 2^{nd} order system; compensatory tracking	analysis of detection threshold; N=5
5. Gai and Curry (1976)	detection threshold on cummulative filter residuals	known model of system and statistical properties of disturbances	step and ramp changes of mean disturbance of 2^{nd} order system; monitoring	reasonable predictions of detection time; N=2
6. Greenstein and Rouse (1982)	detection threshold on likelihood ratio from discriminant function	features independent and linearly weighted	ramp changes of signal to noise ratio of 2^{nd} order system; multiple process monitoring	reasonable predictions of detection time; N=8

operating characteristic or ROC curve. The shape of the curve can be expressed in terms of the human's sensitivity to the signal, while the human's operating point on the curve reflects the aforementioned response threshold.

Signal detection theory has been quite popular with experimental psychologists whose laboratory studies allow rather straightforward manipulations of probabilities, costs, etc. However, in more realistic settings it can be rather difficult to determine the values of these variables and therefore, the model is considerably less useful for realistic situations. Nevertheless, the ROC curve is still a useful way of summarizing human detection performance.

Error vs error rate models

This type of model (see entries 2–4, Table 1) is attractively simple in that human failure detection decisions are assumed to be made solely on the basis of a two-dimensional threshold involving the displayed error and error rate in compensatory tracking tasks (Phatak and Bekey, 1969; Niemela and Krendel, 1975). A related model by Miller and Elkind (1967) assumes that detection is based on the variance of changes in error rate in compensatory tracking. While all three of these models were developed for manual control tasks, the notion of an error vs error rate display is certainly very general and could be applied to other monitoring tasks.

The simplicity of this type of model is not without its disadvantages. In particular, the two-dimensional threshold on error and error rate is highly situation-dependent and the parameters of these models must be empirically adjusted for different types of dynamic processes and failures. Therefore, as with signal detection theory, this type of model is most useful for describing rather than predicting results.

Filter-based models

A Kalman filter is basically a method of resolving the conflict between prediction and subsequent observation (Rouse, 1980). While this conflict may be attributed to poor predictions or noisy observations, an alternative cause of conflict is system failures. Since a failure may change the input–output relationship of a system, predictions based on the pre-failure input–output relationship are likely to disagree with observations of variables produced by the post-failure input–output relationship. The extent of the disagreement can be used as a means of detecting failures.

Gai and Curry (1976) employed this concept to develop a model of human performance in failure detection (see entry 5, Table 1). The sequence of differences between predicted and actual system outputs is obtained from Kalman filter. This sequence of 'residuals' is cumulated and compared with a threshold which is based on acceptable probabilities of missed events and false alarms. Beyond the step and ramp failures noted in Table 1, this model has also been applied to detecting changes in variance and bandwidth (Curry and Govindaraj, 1977).

The strength of this filter-based model is the invariance of its structure across a wide range of dynamic systems. Thus, the model need not be reformulated for each new task situation. There is a cost for this generality, however, in that the model must have explicit knowledge of an appropriate mathematical model of the dynamic process being monitored.

Pattern recognition models

In order to detect failures, humans may observe a wide variety of features beyond errors, error rates, and residuals. For example, sounds, vibrations, and smells may be part of the overall patterns of features relevant to failure detection. In general, the human's detection task is to recognize when the pattern of features is other than normal.

Greenstein and Rouse (1982) have developed a model of human performance in event detection based on discriminant analysis, one of the simplest pattern recognition methods (see entry 6, Table 1). A discriminant function is used to linearly weight any number of task features deemed relevant. This function is used to generate a likelihood ratio which is then compared to a threshold similar to that used in signal detection theory.

The advantages of this model include its ability to overcome a key limitation of signal detection theory (i.e. the process whereby the likelihood ratio is determined), its generalization of the feature-based error vs error rate models, and its ability to function without the mathematical model of the process required for filter-based models. The model's main disadvantages include the need to empirically determine discriminate function coefficients and, when a mathematical model of the process dynamics is available, its inability to make direct use of this information.

Summary

Contrasting the four types of model reviewed here, it seems quite reasonable to conclude that the filter-based models are the best available if the requisite information to use them can be obtained. Otherwise, pattern recognition models are most appropriate. In either case, these models are certainly not available 'off the shelf'. For example, considerable thought would be needed before these models could be applied to resolving the issues associated with the human's relative abilities to detect failures in manually and automatically controlled processes (Ephrath and Young, 1981; Wickens and Kessel, 1981). Nevertheless, sufficient basic research has been performed to justify the effort necessary to apply these models to understanding and resolving these and other issues associated with failure detection.

MODELS OF DIAGNOSIS

Diagnosis refers to the process of identifying the cause of an event. Table 2 summarizes a variety of models of human performance in diagnostic tasks. These models roughly fall into two classes: prescriptive and descriptive. The following discussion elaborates upon this distinction and the summaries given in Table 2.

Prescriptive models

Most of the earlier efforts to model human behavior and performance in fault diagnosis tasks involved comparing human performance to that of prescriptive models. One prescriptive method of diagnosis is the half-split or binary chop which attempts to choose tests that partition the feasible set into two halves in terms of uncertainty. The evidence is fairly conclusive (see entries 2, 3, and 5, Table 2) that humans typically do not make optimal half-split tests (Goldbeck and co-workers, 1957; Dale, 1958; Mills, 1971). While training may help, its effect becomes limited as problem complexity increases. The primary difficulties appear to be humans' inabilities to identify the feasible set and tendencies to utilize irrelevant perceptual cues.

Another type of prescriptive approach to fault diagnosis utilizes probabilities of failure and average action times to find the minimum time solution (Stolurow and co-workers, 1955) or just probabilities to find the most likely fault (Bond and Rigney, 1966). Human's abilities to employ this type of strategy are highly dependent on their knowledge of *a priori* probabilities and average action times (see entires 1 and 4, Table 2). This knowledge is often imperfect, and therefore, humans are precluded from performing as well as the prescribed strategy.

TABLE 2. MODELS OF HUMAN FAILURE DIAGNOSIS

Models \ Attributes	Basic Approach	Key Assumptions	Types of Failure	Experimental Results
1. Stolurow, et al. (1955)	consideration of minimum expected time strategy	repair actions based on probability and time, independent of structure	aircraft powerplants with unequal failure rates and repair times	significant disagreement among instructors about times and probalities; N=10
2. Goldbeck, et al. (1957)	comparison with half-split strategy	diagnosis based on decreasing size of feasible set	components in logic network	significant departure from optimality; difficulty identifying feasible set; N=130
3. Dale (1958)	comparison with half-split strategy	diagnosis based on decreasing size of feasible set	components in flow network	significant departure from optimality; irrelevant cues utilized; N=240
4. Bond and Rigney (1966)	comparison with Bayesian updating of probabilities	test choices based on probability, independent of structure	oscillator circuit	agreement on 50% of solutions; dependent on initial probabilities; N=39
5. Mills (1971)	comparison with half-split strategy	diagnosis based on decreasing size of feasible set	series electrical circuit with unequal failure probabilities	significant departure from optimality; N=6
6. Rasmussen and Jensen (1974)	description of strategies and selection among strategies	verbal protocols reflect strategy	electronic instruments	topographic and symptomatic strategies; strategy fixation; N=6
7. Rouse (1978, 1979)	formation of fuzzy feasible and infeasible sets	diagnosis based on decreasing size of fuzzy feasible set	components in two types of logic network	reasonable predictions of number of tests and effects of aiding; N=36
8. Rouse and Rouse (1979)	information theoretic measure of complexity	strategy of individual affects complexity	components in two types of logic network	reasonable predictions of solution time; N=88
9. Rouse et al. (1980)	rank-ordered set of situation-action rules	rank-ordering of rules fixed	components in two types of logic network	similar choices on 90% of actions; N=154
10. Wohl (1981)	description of possible causes of very skewed repair time distributions	exhaustive search in order of increasing time per action	several military electronics systems	high correlation with average field repair time; N=10 equipments
11. Hunt and Rouse (1982)	fuzzy rank-ordering of pattern recognition and network searching rules	rule choices governed by recall, applicability, usefulness, and simplicity	simulated powerplants and avionics systems	similar choices on 70% of actions; N=10

From this set of comparisons of human performance with that of prescriptive models it can reasonably be concluded that humans are not optimal diagnosticians, at least not with respect to the criteria upon which these models are based. This suboptimality may be due to a lack of knowledge of the prescribed strategy or, due to a lack of knowledge of the requisite information for implementing the strategy or, due to an inability to process the information in the manner required by the strategy. Finally, of course, it could be that humans have performance criteria that include more than just number of actions or time. They may also be concerned with minimizing effort, risk, etc.

Descriptive models

Given that human performance departs substantially from that of prescriptive models, the next logical consideration is describing how humans actually do perform fault diagnosis. While the studies of prescriptive models, particularly the studies of Dale (1958), did attempt to describe deviations from optimality, the studies were not concerned with producing descriptive models *per se*. Such models have only more recently emerged.

The discussion of prescriptive models indicated that humans have difficulty identifying the feasible set. This could be interpreted as meaning that humans find it difficult to crisply say 'yes' or 'no' about the membership of each component in the feasible set of failures. Rouse (1978, 1979) has incorporated this limitation into the half-split strategy by using fuzzy set theory and defining membership in terms of the 'psychological distance' between components and symptoms (see entry 7, Table 2). This model was quite successful in predicting the effects, in terms of number of tests until solution, of providing the human aids for identifying the feasible set.

Some fault diagnosis problems can be solved quickly while others require quite a long time. It seems reasonable to argue that this difference in problem solving time is related to problem complexity. Rouse and Rouse (1979) studied a variety of measures of problem solving complexity and found that an information theoretic measure of the uncertainty associated with all of the connections among components in the feasible set provided the highest correlation between time and complexity (see entry 8, Table 2). Since this measure was based on the feasible set as it evolved through the course of each individual's problem solving, this measure can be said to incorporate each individual's strategy. Thus, the measure reflects the problem solver as well as the problem.

Wohl (1981) has also studied problem solving time as a function of complexity (see entry 10, Table 2). In contrast to the Rouse measure which predicts the time for a specific troubleshooter to solve a particular problem, Wohl's model predicts average repair time across all failures and troubleshooters for a particular piece of equipment. His measure of complexity is based on the number of connections among all components in the piece of equipment. The model also incorporates a particular diagnostic strategy, namely, exhaustive search in order of increasing time per action. The correlation between the average time predictions of this model and data from both field and laboratory studies is quite impressive.

The number of tests and the time until a fault is isolated reflect the *product* of diagnosis. A much better understanding of human problem solving may be possible with models that reflect the *process* of problem solving. Rouse and his colleagues (Rouse and co-workers, 1980) developed a rule-based model that predicted the sequence of actions chosen by the troubleshooter (see entry 9, Table 2). By appropriate choices of rules and rank-orderings, they were able to obtain a high level of agreement between the behavior of the model and that of humans. One particularly interesting conclusion was the fact that better troubleshooters did not necessarily have better rules than poorer troubleshooters; they often simply had a better rank-ordering of the same rules.

Rasmussen's description of diagnostic strategies (Rasmussen and Jensen, 1974; Rasmussen, 1978, 1981) has shown that a variety of strategies are adopted by troubleshooters (see entry 6, Table 2). The most important distinction found was between strategies based on context-specific pattern recognition (i.e. symptomatic strategies) and those based on relatively context-free network searching rules (i.e. topographic strategies). While symptomatic strategies result in a direct mapping from symptoms to hypothesis, topographic strategies involve searching through networks of functional relationships. Rasmussen has also studied the process whereby humans choose and perhaps fixate on strategies.

Hunt and Rouse (see entry 11, Table 2) have combined the fuzzy set model, the rule-based model, and the concepts of Rasmussen to produce a fuzzy rule-based model (Rouse and Hunt, 1981; Hunt and Rouse, 1982). The model has S-rules (symptomatic rules such as: if the car will not start, check the fuel quantity) which it prefers to use if possible, and T-rules (topographic rules such as: if a component's outputs are bad, check its inputs) which it will employ if necessary (i.e. if it cannot find an appropriate S-rule). Particular rules are chosen according to membership in the fuzzy set of choosable rules which is defined as the intersection of the fuzzy sets of recalled, applicable, useful, and simple rules. This model was reasonably successful in predicting the sequence of actions chosen by aircraft mechanics in troubleshooting simulated powerplant and avionics systems. Further, it was useful for illustrating the shift from S-rules to T-rules when unfamiliar problems were encountered.

Summary

Contrasting the variety of models of diagnostic behavior presented in this section, it is clear that the unconstrained prescriptive models do not provide good descriptions of human behavior and performance. For the purpose of predicting repair time for a particular equipment system, averaged across types of failure and different troubleshooters, Wohl's model is probably the best choice. For more fine-grained predictions of human behavior, the model of Hunt and Rouse, as a derivative of earlier work by Rasmussen and Rouse, would seem to offer the most appropriate approach.

MODELS OF COMPENSATION

If a failure must be diagnosed during system operation, as opposed to during maintenance, then the human problem solver typically must be concerned with both keeping the system operating and diagnosing the source of the problem. The process of sustaining system operation in failure situations is termed compensation. In this section, two types of compensation will be considered: (1) compensating for symptoms and (2) compensating for failures.

Compensating for symptoms

There appear to be two general types of symptom. The first type includes abnormal and emergency events such as fires, leaks, system trips, etc. These events are often dealt with using standard procedures, which may or may not be formalized. In some situations, however, there are no written or unwritten procedures and humans must revert to problem solving.

The second type of symptom includes substantial deviations of state variables. Examples include pressures, levels, flows, and velocities that are too high or too low. This type of symptom can be dealt with in several ways. For loops that are normally automatically controlled, the operator may compensate by assuming manual control. Another approach to compensation is through reallocation of resources including switching to backup modes, re-routing of resources, and re-prioritizing or shedding of demands on the system. With this second approach, standard procedures are typically not available and problem solving is required.

Compensating for failures

Once the diagnostic process has proceeded to the point of having identified the failure, the operator must decide how to compensate for the failed component. The most obvious compensation is to repair or replace the failed component. However, this is often not immediately possible. As a result, the operator may have to plan for degraded mode operation.

Degraded mode operation involves continued operation with the loss or substantial impairment of one or more system functions. If the lost or impaired functions are critical (e.g. engines to an aircraft), then the only thing of interest may be that of bringing system operation to a halt in a safe and orderly manner (e.g. a safe landing). For critical situations such as these there are often standard procedures.

Many situations, however, do not call for or allow suspending of systems operations. In these cases, the operator may have to continue compensating for the symptoms and/or plan for operating with a long-term loss or impairment of functions. This may require problem solving.

Coordinating compensation and diagnosis

Compensation and diagnosis can be viewed as two separate tasks competing for the operator's attention (Rouse and Morris, 1981a, b; van Eekhout and Rouse, 1981). Unfortunately, this situation can result in the operator focusing on one task to the exclusion of the other. If the effects of failures are cumulative or if symptoms tend to become aggravated, focusing on only compensation or diagnosis can have disasterous consequences. For example, at least one major airline crash has been attributed to the crew having focused on diagnosis to the exclusion of all other tasks.

The coordination of compensation and diagnosis is not simply the process of managing two tasks simultaneously. Since the two tasks are far from independent, they have the potential for being conflicting. For example, compensating for the symptoms may make diagnosis more difficult. On the other hand, the two tasks may be complementary in that information acquired for performing one task may provide information valuable to performing the other task. In either case (i.e. conflicting or complementary), this interdependence illustrates the potential complexity of dealing with problem solving at multiple levels.

Models

Unlike the discussion of detection and diagnosis, this section has no tabulation of models and their attributes. For those situations where compensation involves executing standard procedures or manual control, there are a variety of models available (Sheridan and Ferrell, 1974; Rouse, 1980; 1981a). However, these are not models of human problem solving. In fact, other than the aforementioned models by Rasmussen, Rouse, and Hunt, there are no directly applicable models of problem solving behavior in coordinating compensation and diagnosis, or for compensation itself. The next section proposes an outline for such a model.

OUTLINE FOR AN OVERALL MODEL

From the foregoing discussion, it is obvious that considerable effort has been invested in the study of human problem solving in general, and human detection, diagnosis, and compensation for system failures in particular. However, most of the models discussed thus far focus on a single aspect of problem solving. Only a few of the models (Newell and Simon, 1972; Minsky, 1975; Schank and Abelson, 1977; Rasmussen, 1979) consider the full breadth and robustness of human problem solving behavior. What is needed is a model that, at least conceptually, captures the whole of problem solving and, at the same time, can be operationalized within specific task domains. This section outlines the specifications for such a model.

Pattern recognition orientation

A conclusion that surfaced repeatedly in the earlier discussions was that humans, if given a choice, would prefer to act as context-specific pattern recognizers rather than attempting to calculate or optimize. Obviously, life would be difficult indeed if one had to constantly recalculate various things in order to make choices. Thus, human preference for pattern recognition is justifiable both scientifically and practically.

However, if the human does not recognize a pattern, a mode of problem solving other than context-specific pattern recognition must be employed. The alternative modes may be called heuristic, analytical, topographic, etc. A common characteristic of these modes is that the human must go beyond the surface features of the problem. Since the focus of this paper is on system failures, this notion can be more precisely stated as the human must go beyond the system state and consider the system structure.

Figure 1 illustrates this fundamental concept. The human is assumed to have a clear preference for proceeding on the basis of state information. The use of structural information is definitely a less-preferred alternative. The mechanism shown in Fig. 1, which is elaborated upon throughout the remainder of this section, is proposed as the central and only mechanism necessary for an overall model.

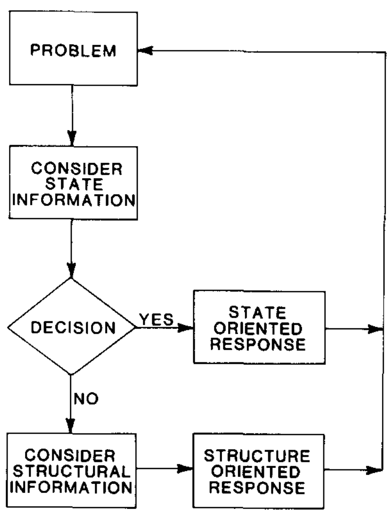

FIG. 1. Basic mechanism for proposed model of human problem solving.

Levels of problem solving

Considering the literature reviewed earlier in this paper and a variety of studies of human problem solving in aircraft, ships, and process plants (Johannsen and Rouse, 1980, 1981; Rouse and co-workers, 1982; Johnson and Rouse, 1982a, b; van Eekhout and Rouse, 1981; Rouse and Morris, 1981a, b), it seems reasonable to conclude that problem solving occurs on several levels. Perhaps the most obvious example of multilevel problem solving is the aforementioned coordination of compensation and diagnosis. The concept of multiple levels is, however, much more general than the idea of coordinating tasks.

It appears that three general levels of problem solving are needed to model human behavior: (1) recognition and classification, (2) planning, and (3) execution and monitoring. Recognition and classification involves detecting that a problem solving situation exists and assigning it to a category. Planning is the process whereby the approach to solving a problem is determined. Execution and monitoring is the actual process of solving the problem.

Table 3 summarizes how the basic mechanism in Fig. 1 applies to the three levels of problem solving. At the highest level (i.e. recognition and classification), the human is assumed to identify the context and category of a problem. If the human finds the state information to match an available frame (Minsky, 1975), problem solving proceeds on that basis. If an appropriate frame is not in the human's repertoire, structural information might provide clues to an analogy or be used to employ basic principles of, for example, the scientific method.

TABLE 3. DECISIONS AND RESPONSES FOR THREE LEVELS OF PROBLEM SOLVING

Level \ Process	Decision	State-Oriented Response	Structure-Oriented Response
1. Recognition and Classification	Frame Available?	Invoke Frame	Use Analogy and/or Basic Principles
2. Planning	Script Available?	Invoke Script	Formulate Plan
3. Execution and Monitoring	Pattern Familiar?	Apply Appropriate S-Rule	Apply Appropriate T-Rule

At the next level (i.e. planning), the human must decide how the problem will be attacked. Based on the state information, the human may conclude that the problem solving situation is familiar and the appropriate script (Schank and Abelson, 1977) or standard procedure can be employed. If no script is available, the human must use structural information to plan in terms of generating alternatives, imagining consequences, valuing consequences, and so on (Johannsen and Rouse, 1979).

Actual problem solving occurs at the lowest level (i.e. execution and monitoring) where scripts or plans are executed and monitored for success. Familiar patterns of state information may allow for the use of context-specific symptomatic rules (S-rules) that map directly from observation to hypothesis or action (e.g. if the engine will not crank, check the battery). If the pattern is not familiar, structural information may allow the use of topographic rules (T-rules) for searching the structure of the problem (e.g. if a component's inputs are good and its outputs bad, the component has failed).

All of the responses noted in Table 3 (i.e. invoke frame, use analogy, etc.) invoke the same mechanism as shown in Fig. 1. This mechanism is recursively invoked until actions are produced and the problem solved. Thus, in contrast to many of the models discussed earlier in this paper, the model outlined here is very simple, involving a single mechanism that is recursively employed for all aspects of problem solving.

Hierarchical vs heterarchical

If one views problem solving from an operations research or management science perspective, one should hierarchically consider goals, objectives, attributes, alternatives, etc. This hierarchical approach has often been adopted by computer scientists when designing knowledge-based expert problem solving systems (Sacerdoti, 1974). The model proposed in this paper certainly could perform hierarchically by first invoking a frame or an analogy, then invoking a script or planning, and finally acting via S-rules and T-rules.

However, it has been argued that human problem solving is heterarchical or opportunistic rather than hierarchical (Hayes-Roth and Hayes-Roth, 1979). In other words, the human does not solve problems in purely a top-down or bottom-up manner. Instead, it appears that the human operates on all levels almost simultaneously.

The proposed model can produce this type of behavior if one assumes that the three decisions (i.e. frame available, script available, and pattern familiar) are constantly, but not necessarily consciously, being re-evaluated. Thus, for example, the model might be using T-rules to plan on the basis of an analogy and suddenly realize the applicability of a script. This could result in the preemption of planning and the rapid application of a sequence of S-rules which provides new information and results in a familiar frame being recognized which leads to a new script and so on.

If all three decisions are constantly being re-evaluated, it is possible that conflicts will arise in terms of which decision should take precedence. Such conflicts might be resolved by giving more credence to closer matches (e.g. a very familiar pattern is more captivating than a somewhat familiar script). Another method of resolving conflict is to place more weight on alternatives that maintain the current direction of the problem solving. In other words, the model could incorporate the assumption that the human would like to avoid changes in frame, script, etc.

Behaviour of the model

Since the model outlined here represents a synthesis of the many concepts and models reviewed in earlier sections of this paper, it should not be surprising that this model will produce the types of behavior noted in those discussions. The strength of the model is that a very simple mechanism and method of organization can produce such an impressive range of behaviors. Of course, this possibility, from a slightly different perspective, has been investigated by others (Newell and Simon, 1972).

A particularly interesting aspect of the model's behavior, as well as that of humans, is its potential for making errors. The model has two inherent possibilities for causing errors. The first possibility relates to the model's recursive use of the basic mechanism in Fig. 1. As the model recursively invokes this mechanism, it needs a 'stack' or some short-term memory for keeping track of where it is and how it got there. If short-term memory is limited, as it is in humans, the model may recurse its way into getting lost or, pursuing tangents from which it never returns. To constrain this phenomenon, it is probably reasonable to assume that lower priority items in the stack are more likely to be lost first. For example, one is more likely to forget one's umbrella than to forget to go to work.

The second possibility for causing errors is the matching of irrelevant or inappropriate patterns. For example, the model, or a human, may be captured by an inappropriate but similar script or S-rule. As a result, the model may pursue an inappropriate path until it suddenly realizes, perhaps much too late to be able to recoup, that it has wandered far afield from where it thought it was headed.

The fact that the proposed model has inherent possibilities for making errors, particularly somewhat subtle errors, provides an interesting avenue for evaluating the model. Most models are evaluated in terms of their abilities to achieve the same levels of desired task performance as humans. A much stronger test would involve determining if the model deviates from desired performance in the same way and for the same reasons as humans. The proposed model can potentially be evaluated in this manner.

Summary

In this section, an outline of a model of human problem solving has been proposed. This outline is, in a sense, a set of model specifications based on a synthesis of a wide range of concepts and models as well as a variety of experimental results. The strength of the proposed model is its potential ability to represent a wide range of human problem solving behavior while also being readily implementable for evaluation.

CONCLUSIONS

Considering the relationship of the overall model outlined here to the other models discussed in this paper, this conceptual model is, to a great extent, an outgrowth of earlier work by Rouse and Hunt, and also by Rasmussen. The frame, script, and S-rule aspects of the model are also fairly similar to concepts proposed by Newell and Simon (1972). The manner in which the model utilizes both state and structural information, and the recursive use of the same basic mechanism on all levels of problem solving are perhaps the model's most unique characteristics. In this way, the model is potentially capable of dealing with unfamiliar problem contexts via analogies, T-rules, etc.

Comparing the proposed model with the many models of detection and diagnosis reviewed in this paper, the pattern recognition oriented and rule-based models can be easily incorporated within the general framework outlined here. The models which assume the human to perform some type of calculation (e.g. Bayesian, filter-based, and fuzzy half-split models) best fit within the specific portions of this framework for planning and T-rules. It should be noted that the assumptions underlying these calculation oriented models differ substantially from those of the proposed model which assumes the human to avoid calculation if at all possible.

The model's ability to operate almost simultaneously on several levels provides the potential for representing the coordination of compensation and diagnosis. Because of short-term memory limitations, the model also allows for the possibility of errors in coordination (e.g. focusing on one task to the exclusion of the other). What is not clear at this point, and is the topic of several current investigations, is the nature of S-rules, T-rules, scripts, and plans relative to coordinating compensation and diagnosis. This important topic deserves considerable study.

The basic premise of this paper is that the responsibilities of the human operator will increasingly be dominated by problem solving. The design of systems to support the human operator in fulfilling these responsibilities should be based on knowledge of human abilities and limitations in problem solving. This paper has reviewed the state of the art in this area and proposed specifications for a new model which has the potential for being a vehicle for integrating and advancing understanding of human problem solving.

Acknowledgements—This research was supported in part by the U.S. Army Research Institute for the Behavioral and Social Sciences under contract no. MDA 903-79-C-0421 and in part by the National Aeronautics and Space Administration under Ames grant no. NAG 2-123.

REFERENCES

Albus, J. S. (1970). Mechanisms of planning and problem solving in the brain. *Math. Biosci.*, **45**, 247.

Atwood, M. E. and P. G. Polson (1976). A process model for the water jug problem. *Cog. Psych.*, **8**, 191.

Atwood, M. E., P. G. Polson, R. Jefferies and H. R. Ramsey (1978). *Planning as a Process of Synthesis.* Science Applications, Englewood, CO. Report SAI-78-144-DEN.

Bond, N. A. and J. W. Rigney (1966). Bayesian aspects of trouble-shooting behavior. *Hum. Fac.*, **8**, 377.

Bree, D. S. (1975). Understanding of structured problem solutions. *Instruc. Sci.*, **3**, 327.

Brooke, J. B. and K. D. Duncan (1980). An experimental study of flowcharts as an aid to identification of procedural faults. *Ergonomics*, **23**, 387.

Brooke, J. B., and K. D. Duncan (1981). Effects of system display format on performance in a fault location task. *Ergonomics*, **24**, 175.

Brooke, J. B., K. D. Duncan and C. Cooper (1980). Interactive instruction in solving fault finding problems—an experimental study. *Int. J. Man–Mach. Stud.*, **12**, 217.

Brown, J. S., R. R. Burton and A. G. Bell (1975). SOPHIE: a step toward creating a reactive learning environment. *Int. J. Man–Mach. Stud.*, **7**, 675.

Burroughs, M. M. (1979). An investigation of mental coding mechanisms and heuristics used in electronics trouble-shooting. University of Oklahoma, Engng. D. thesis.

Chase, W. G. and H. A. Simon (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual Information Processing.* Academic Press, New York.

Curry, R. E. and T. Govindaraj (1977). The human as a detector of changes in variance and bandwidth. *Proc. Thirteenth Ann. Conf. on Manual Control*, Cambridge, MA, pp. 217–221.

Dale, H. C. A. (1958). Fault finding in electronic equipment. *Ergonomics*, **1**, 356.

Dreyfus, H. L. and S. E. Dreyfus (1979). The psychic boom: flying beyond the thought barrier. Operations Research Center, Berkeley, CA, Report ORC 79-3.

Duncan, K. D., and M. J. Gray (1975). An evaluation of a fault finding training course for refinery operators. *J. Occup. Psych.*, **48**, 199.

Duncan, K. D. and A. Shepherd (1975). A simulator and training technique for diagnosing plant failures from control panels. *Ergonomics*, **18**, 627.

Ephrath, A. R. and L. R. Young (1981). Monitoring vs. man-in-the-loop detection of aircraft control failures. In J. Rasmussen and W. B. Rouse (Eds), *Human Detection and Diagnosis of System Failures.* Plenum Press, New York, pp. 143–154.

Freedy, A. and L. F. Lucaccini (1981). Adaptive computer training system (ACTS) for fault diagnosis in maintenance tasks. In J. Rasmussen and W. B. Rouse (Eds.), *Human Detection and Diagnosis of System Failures.* Plenum Press, New York, pp. 637–658.

Gai, E. G. and R. E. Curry (1976). A model of the human observer in failure detection tasks. *IEEE Trans Syst., Man & Cybern.*, **SMC-6**, 85.

Gerwin, D. (1974). Information processing, data inferences, and scientific generalization. *Behav. Sci.*, **19**, 314.

Gerwin, D. and P. Newstad (1977). A comparison of some inductive inference models. *Behav. Sci.*, **22**, 1.

Glaser, R. D., E. Damrin and F. M. Gardner (1954). The tab item: a technique for the measurement of proficiency in diagnostic problem solving tasks. *Educ. and Psych. Meas.*, **14**, 283.

Goldbeck, R. A., B. B. Bernstein, W. A. Hillix and M. H. Marx (1957). Application of the half-split technique to problem-solving tasks. *J. Exp. Psych.*, **53**, 330.

Greenstein, J. S. and W. B. Rouse (1982). A model of human decision making in multiple process monitoring situations. *IEEE Trans Syst., Man & Cybern.*, **SMC-12**, 182.

Gregg, L. W. and H. A. Simon (1967). Process models and stochastic models of simple concept formulation. *J. Exp. Psych.*, **4**, 246.

Hayes, J. R. (1973). On the function of visual imagery in elementary mathematics. In W. G. Chase (Ed.), *Visual Information Processing*. Academic Press, New York.

Hayes-Roth, B. and F. Hayes-Roth (1979). A cognitive model of planning. *Cog. Sci.*, **3**, 275.

Henneman, R. L. and W. B. Rouse (1982). Measures of human performance in fault diagnosis tasks, to be published.

Hormann, A. M. (1971). A man–machine synergistic approach to planning and creative problem solving: I. *Int. J. Man-Mach. Stud.*, **3**, 167.

Huesman, L. R. and C.-M. Cheng (1973). A theory for induction of mathematical functions. *Psych. Rev.*, **80**, 126.

Hunt, R. M. and W. B. Rouse (1981). Problem solving skills of maintenance trainees in diagnosing faults in simulated powerplants. *Hum. Fac.*, **23**, 317.

Hunt, R. M. and W. B. Rouse (1982). A fuzzy rule-based model of human problem solving. *Proc. Am. Control Conf.*, Washington.

Jacoby, L. L. (1978). On interpreting the effects of repetition: solving a problem versus remembering a solution. *J. Verb. Learn. and Verb. Behav.*, **17**, 649.

Jeffries, R., P. G. Polson, L. Razran and M. E. Atwood (1977). A process model for missionaries–cannibals and other river-crossing problems. *Cog. Psych.*, **9**, 412.

Johannsen, G. and W. B. Rouse (1979). Mathematical concepts for modeling human behavior in complex and man–machine systems. *Hum. Fac.*, **21**, 733.

Johannsen, G. and W. B. Rouse (1980). A study of the planning process of aircraft pilots in emergency and abnormal situations. *Proc.* 1980 *IEEE Int. Conf. on Cybernetics and Society*, Boston, pp. 738–745.

Johnannsen, G., and W. B. Rouse (1981). Problem solving behavior of pilots in abnormal and emergency situations. *Proc. First Ann. European Conf. on Human Decision Making and Manual Control*, Delft, The Netherlands, pp. 142–150.

Johnson, E. S. (1978). Validation of concept learning strategies. *J. Exp. Psych.*, **107**, 237.

Johnson, S. L. (1981). Effect of training device on retention and transfer of a procedural task. *Hum. Fac.*, **23**, 257.

Johnson, W. B. and W. B. Rouse (1982a). Analysis and classification of human errors in troubleshooting live aircraft powerplants. *IEEE Trans Syst., Man & Cybern.*, **SMC-12**, 389.

Johnson, W. B. and W. B. Rouse (1982b). Training maintenance technicians for troubleshooting: two experiments with computer simulations. *Hum. Fac.*, **24**, 271.

Kearsley, G. P. (1975). Problem-solving set and functional fixedness: a conceptual dependency hypothesis. *Canad. Psych. Rev.*, **16**, 261.

Kozielecki, J. (1972). A model for diagnostic problem solving. *Acta Psych.*, **36**, 370.

Landa, S. (1972). CATTS: computer-aided training in troubleshooting, Rand Corporation, Santa Monica, CA. Report R-518-PR.

Loftus, E. F. and P. Suppes (1972). Structural variables that determine problem-solving difficulty in computer-assisted instruction. *J. Ed. Psych.*, **63**, 531.

Luger, G. F. and M. A. Bauer (1978). Transfer effects in isomorphic problem situations. *Acta Psych.* **42**, 121.

Malin, J. T. (1979). Information processing load in problem solving by network search. *J. Exp. Psych.*, **5**, 379.

Mason, E. J., W. J. Bramble and T. A. Mast (1975). Familiarity with context and conditional reasoning. *J. Ed. Psych.*, **67**, 238.

Mayer, R. E. (1975). Different problem-solving competencies established in learning computer programming with and without meaningful models. *J. Ed. Psych.*, **67**, 725.

Miller, D. C. and J. I. Elkind (1967). The adaptive responses of the human controller to sudden changes in controlled process dynamics. *IEEE Trans Hum. Factors in Electronics*, **HFE-8**, 218.

Mills, R. G. (1971). Probability processing and diagnostic search: 20 alternatives, 500 trials. *Psychonomic Sci.*, **24**, 289.

Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The Psychology of Computer Vision*. McGraw-Hill, New York.

Neimark, E. D. and J. L. Santa (1975). Thinking and concept attainment. *Ann. Rev. Psych.*, **26**, 173.

Nuebauer, H. N. and W. B. Rouse (1979). A study of the use of color graphics for representing schematic information. *Proc. of the Fifteenth Ann. Conf. on Man. Control.* Wright State University, pp. 624–631.

Newell, A. and H. A. Simon (1972). *Human Problem Solving*. Prentice-Hall, Englewood Cliffs, NJ.

Niemela, R. J. and E. S. Krendel (1975). Detection of a change in plant dynamics in a man–machine system. *IEEE Trans Syst., Man & Cybern.*, **SMC-5**, 615.

Nievergelt, J. (1977). Information content of chess positions. *SIGART Newsletter*, **62**, 13.

Peters, J. T., K. R. Hammond and D. A. Summers (1974). A note on intuitive vs. analytical thinking. *Organ. Behav. and Hum. Perform.*, **12**, 125.

Phatak, A. V. and G. A. Bekey (1969). Decision processes in the adaptive behavior of human controllers. *IEEE Trans Syst., Sci. & Cybern.*, **SSC-5**, 333.

Polich, J. M. and S. H. Schwartz (1974). The effect of problem size on representation in deductive problem solving. *Mem. and Cog.*, **2**, 683.

Rasmussen, J. (1978). Notes on diagnostic strategies in process plant environment. Risø National Laboratory, Roskilde, Denmark, Report M-1983.

Rasmussen, J. (1979). On the structure of knowledge. A morphology of mental modes in a man–machine context. Risø National Laboratory, Roskilde, Denmark, Report M-2192.

Rasmussen, J. (1981). Models of mental strategies in process plant diagnosis. In J. Rasmussen and W. B. Rouse (Eds), *Human Detection and Diagnosis of System Failures*. Plenum Press, New York, pp. 241–258.

Rasmussen, J. and A. Jensen (1974). Mental procedures in real life tasks: a case study of electronic troubleshooting. *Ergonomics*, **17**, 293.

Rasmussen, J. and W. B. Rouse (Eds) (1981). *Human Detection and Diagnosis of System Failures*. Plenum Press, New York.

Rouse, S. H., W. B. Rouse and J. M. Hammer (1982). Design and evaluation of an onboard computer-based information system for aircraft. *IEEE Trans Syst., Man & Cybern.*, **SMC-12**, 451.

Rouse, W. B. (1978). A model of human decision making in a fault diagnosis task. *IEEE Trans Syst., Man & Cybern.*, **SMC-8**, 357.

Rouse, W. B. (1979). A model of human decision making in fault diagnosis tasks that include feedback and redundancy. *IEEE Trans Syst., Man & Cybern.*, **SMC-9**, 237.

Rouse, W. B. (1980). *Systems Engineering Models of Human–Machine Interaction*. North-Holland, New York.

Rouse, W. B. (1981a). Human-computer interaction in the control of dynamic systems. *Computing Survey*, **12**, 71.

Rouse, W. B. (1981b). Experimental studies and mathematical models of human problem solving performance in fault diagnosis tasks. In J. Rasmussen and W. B. Rouse (Eds), *Human Detection and Diagnosis of System Failures*. Plenum Press, New York, pp. 199–216.

Rouse, W. B. (1982). A mixed-fidelity approach to technical training. *J. Ed. Tech. Sys.*, **11**.

Rouse, W. B. and R. M. Hunt (1981). A fuzzy rule-based model of human problem solving in fault diagnosis tasks. *Proc. Eighth World Congress of IFAC*, Kyoto, pp. 95–101.

Rouse, W. B. and N. M. Morris (1981a). Human problem solving in a process control task. *Proc. First Ann. European Conf. of Human Decision Making and Manual Control*, Delft, The Netherlands, pp. 218–226.

Rouse, W. B. and N. M. Morris (1981b). Studies of human problem solving in a process control task. *Proc.* 1981 *IEEE Int. Conf. on Cybernetics and Society*, Atlanta, pp. 479–487.

Rouse, W. B. and S. H. Rouse (1979). Measures of complexity of fault diagnosis tasks. *IEEE Trans Syst., Man & Cybern.*, **SMC-9**, 720.

Rouse, W. B., S. H. Rouse and S. J. Pellegrino (1980). A rule-based model of human problem solving performance in fault diagnosis tasks. *IEEE Trans Syst., Man & Cybern.*, **SMC-10**, 366.

Rumelhart, D. E. and A. A. Abrahamson (1973). A model of analogical reasoning. *Cog. Psych.* **5**, 1.

Sacerdoti, E. D. (1974). Planning in a hierarchy of abstraction spaces. *Art. Intell.*, **5**, 115.

Schank, R. C. and R. P. Abelson (1977). *Scripts, Plans, Goals, and Understanding.* Lawrence Erlbaum, Hillsdale, NJ.

Shepard, A., E. C. Marshall, A. Turner and K. D. Duncan (1977). Diagnosis of plant failures from a control panel: a comparison of three training methods. *Ergonomics*, **20**, 347.

Sheridan, T. B. and W. R. Ferrell (1974). *Man–Machine Systems: Information, Control, and Decision Models of Human Performance.* M.I.T. Press, Cambridge, MA.

Sheridan, T. B. and G. Johannsen (Eds) (1976). *Monitoring Behavior and Supervisory Control.* Plenum Press, New York.

Siegler, R. S. (1977). The twenty questions game as a form of problem solving. *Child Dev.*, **48**, 395.

Simon, H. A. (1975). The functional equivalence of problem solving skills. *Cog. Psych.*, **7**, 268.

Simon, H. A. and S. K. Reed (1976). Modeling strategy shifts in a problem solving task. *Cog. Psych.*, **8**, 86.

Simonton, D. K. (1975). Creativity, task complexity, and intuitive versus analytical problem solving. *Psych. Rep.*, **37**, 351.

Smith, J. P. (1973). The effect of general versus specific heuristics in mathematical problem solving, Columbia University, Ph.D. thesis.

Sternberg, R. J. (1977). Component processes in analogical reasoning. *Psych. Rev.*, **84**, 353.

Sternberg, R. J. and E. M. Weil (1980). An aptitude–strategy interaction in linear syllogistic reasoning. *J. Ed. Psych.*, **72**, 226.

Stolurow, L. M., B. Bergum, T. Hodgson and J. Silva (1955). The efficient course of troubleshooting as a joint function of probability and cost. *Educ. & Psych. Meas.*, **15**, 463.

Towne, D. M. (1981). A general purpose system for simulating and training complex diagnosis and troubleshooting tasks. In J. Rasmussen and W. B. Rouse (Eds), *Human Detection and Diagnosis of System Failures.* Plenum Press, New York, pp. 621–635.

van Eekhout, J. M. and W. B. Rouse (1981). Human errors in detection, diagnosis, and compensation for failures in the engine control room of a supertanker. *IEEE Trans Syst., Man & Cybern.*, **SMC-11**, 813.

Wason, P. C. and P. N. Johnson-Laird (1972). *The Psychology of Reasoning.* Harvard University Press, Cambridge, MA.

Waterman, D. A. and F. Hayes-Roth (1978). (Eds.), *Pattern-Directed Inference Systems.* Academic Press, New York.

Wickens, C. D. and C. Kessel (1981). Failure detection in dynamic systems. In J. Rasmussen and W. B. Rouse (Eds), *Human Detection and Diagnosis of System Failures.* Plenum Press, New York, pp. 155–169.

Wohl, J. G. (1981). System complexity, diagnostic behavior, and repair time: a predictive theory. In J. Rasmussen and W. B. Rouse (Eds.), *Human Detection and Diagnosis of System Failures.* Plenum Press, New York, pp. 217–230.

Wood, D., J. Shotter and D. Godden (1974). An investigation of the relationships between problem solving strategies, representation, and memory. *Quart. J. Exp. Psych.*, **26**, 252.