

DRA/AMES

NASA Grant NAG 2-123\*

PILOT INTERACTION WITH AUTOMATED AIRBORNE DECISION MAKING SYSTEMS

Semiannual Progress Report

March 1984 - August 1984

William B. Rouse, Principal Investigator

John M. Hammer

Nancy M. Morris

Annette E. Knaeuper

Edward N. Brown

Charles M. Lewis

Wan C. Yoon

(NASA-CR-177002)	PILOT INTERACTION WITH	N86-295C8
	AUTOMATED AIRBORNE DECISION MAKING SYSTEMS	
	Semiannual Progress Report, Mar. - Aug. 1984	
(Georgia Inst. of Tech.)	45 p	CSCL 05H
		Unclas
		G3/54 43298

Center for Man-Machine Systems Research  
 Georgia Institute of Technology  
 Atlanta, Georgia 30332

\*The NASA Technical Officer for this grant is Everett Palmer, NASA Ames Research Center, Moffett Field, CA 94035.

## INTRODUCTION

Increased requirements for safety and efficiency as well as increased availability of reliable and inexpensive computer technology have resulted in a trend of more and more computers being employed in flight management. However, this trend by no means indicates that human operators will disappear from aircraft cockpits. Instead, it means that the roles of the pilot, copilot, and flight engineer will evolve to include increased responsibilities for monitoring and supervising the various computer-based systems employed in the aircraft.

While this assessment of the future roles of the members of the flight crew is fairly easy to accept, it is certainly not straightforward to decide how various flight tasks should be allocated among humans and computers. Further, it is not clear how humans and computers should communicate regarding the process by which their tasks are performed and the products that result. This report discusses progress of a research program whose overall objectives include providing at least partial answers to some of the questions surrounding these issues.

The following two sections discuss two project areas which are currently being pursued in this program of research: 1) the intelligent cockpit and 2) studies of human problem solving. The first area involves an investigation of the use of advanced software engineering methods (e.g., from artificial intelligence) to aid aircraft crews in procedure selection and execution. The second area is focused on human problem solving in dynamic environments, particularly in terms of identification of rule-based models and alternative approaches to training and aiding. Both of these efforts are producing results that are planned to be tested further in the Center's evolving full-scope flight simulation facility. Progress on developing this facility is discussed in the section on the intelligent cockpit.

## THE INTELLIGENT COCKPIT

Design goals for the intelligent aid are in the following section. A review of progress in developing the DC8 flight simulator is in the second section.

### GOALS FOR ANALYSIS OF DATA IN THE INTELLIGENT COCKPIT PROGRAM

#### The Problem

The data are a large, rich state vector of the aircraft. The intelligent aid is to monitor this data to watch for pilot checklist error and unsafe conditions either present or future.

The current approach to monitoring has been to divide the flight into phases according to rules that examine the data past and present. It is possible to be more specific about a phase than about the whole flight. Thus, it is possible to apply pre-programmed checks as appropriate to that phase. While this approach has merits, there are design limits to what can be preprogrammed. In other words, this approach works only for the situations the designer anticipates. In particular, this approach can be made to look quite good in a controlled experiment where the intelligent aid has been programmed to aid on those situations the pilot will encounter. Our approach to demonstrating intelligent aid concepts is to build something that at face value will handle a wide variety of situations. If, due to complexity, a complete aid cannot be constructed, we would prefer an in-depth aid for a particular problem (e.g., electrical malfunctions). We reject an aid that is able to catch a few problems of all kinds but which appears inadequate for complete coverage.

## External Goals

The external goals are the visible product that an observer or pilot would see when the program is running. The following are planned.

1. The program will drive two displays. The first will be the experimenter - system designer display. It will be used for program debugging, etc. The second will be a display for the pilot. It will be created even if there is no pilot to observe it (e.g., off line simulator data). This generality will allow the aid to be incorporated into the DC8 flight simulator without substantial reworking. It will also allow subjective pilot evaluation before being placed in the simulator.
2. The pilot's display will feature the following. Procedures will automatically be selected for display by the aid. When the aid detects an error, the procedure display will be changed, perhaps by highlighting.
3. The display will alert the pilot to aircraft operation outside the normal regime at the present time and in the future.

## Internal Goals

The internal goals describe how the program is organized. Some of these supplement the external goals; some are not apparent from that vantage point.

1. The program will understand the procedure in terms of a model of the aircraft. This model will be used to predict the future states of the aircraft. The program will understand the effect of procedural errors, not just know how to detect them as was the case in the earlier version of the program.

2. To the maximum extent possible, the program will generally apply to all commercial transport aircraft. This will facilitate changing from NASA's B727 to our DC8.
3. Existing scripts will be enhanced and will represent more than a procedure checklist. The script will be active whenever the aircraft is in the appropriate state. While the aircraft is in the script, constraints on safe flight will be constantly monitored.
4. The program's goal will be to avoid aborting the flight plan. This goal is made up of a number of subgoals which are by aircraft subsystem. For example, to avoid aborting the flight plan, the following failures must be avoided

- propulsion failure
- aerodynamic failures
- hydraulic failures
- etc.

Each of these areas can in turn be broken down into finer problems

- propulsion failure
  - over stress engines
  - engine ice up
  - engine stall
  - engine oil pressure
  - engine explosion and airframe/airfoil damage
  - etc.

To be systematic, it would be important to enumerate as many forms of failures as possible. Organizing them by subsystems would tend to keep the designer from missing a failure mode.

To restrict the aid program to a manageable size, it may be necessary to aid only a subset of the potential failures (e.g., only electrical problems). This would, however, demonstrate the

approach's generality, and it would be clear how to generalize to the entire aircraft.

5. The program will be coded as a rule-based system in LISP. As we have stated before, the problem of intelligent aiding is principally a symbolic problem. In addition, the debugging tools for LISP are far superior to those of other languages. Also, there is a practical advantage to the use of LISP. If the simulator code or the intelligent aid should end up consuming too much of the VAX's capability, it should be possible to move the LISP aid code to a LISP machine. This should be considerably cheaper than other VAX.
6. The program must make some predictions about the aircraft's future. This is necessary to check for safe, future states. Representing and manipulating the passage of time is a recognized problem in artificial intelligence. Hopefully, some contribution can be made to this area.

### Conclusion

The proposed new aid will have an improved understanding of the aircraft and consequently the ability to alert the pilot to present and future dangers. The aid will also be useable on both the B727 and the DC8.

### SIMULATOR DEVELOPMENT

#### Simulator Hardware

Most of the switches and controls (overhead panels and flight controls) have wires connected that go outside the cockpit and are waiting to be connected to the A/D converter. A few engine controls and

the elevator controls must have sensors designed and installed before being wired.

All of the CRT displays have been mounted in the flight deck. A cowling must be added to enclose one of the CRT's forward of the pedestal. Both this cowling and a force feel system for the control yoke are being designed and built in the GTRI shop. The custom keyboards have been designed and are currently being fabricated. A force feel system for the pedals is installed and undergoing final adjustment.

#### Simulation Software

The simulation software currently supports high fidelity engines, flight dynamics, radio navigation and an autopilot. Hydraulic, electrical, and fuel systems are not being simulated in any but the most elementary ways in this version of the simulator. The remaining work on the simulation software is to integrate it with the display software and to modify the takeoff routines to use flight dynamics. Some minor modifications (coefficient changes and term changes) are being made to the flight dynamics to make it perform as a commercial transport.

#### Display Software

There are three displays to be produced for in the simulator: the flight instruments, the engine status, and the navigation/autopilot/communication display. The engine status display is being programmed now; it is fairly simple. The navigation/autopilot/communication display specifications have been worked out and will be programmed after the engine display is completed.

The flight instrument display is by far the most complicated display. The original plans were to use an Apple II to drive the color

displays. It appears this will not work because the Apple is not fast enough. The following simple analysis shows how this was determined. The ADI must be updated frequently; at a minimum, updating this display requires drawing of 11 or 12 lines. Timing estimates show the Apple (using Pascal) can erase an old line and draw a new line in  $.02 + .0002 * \text{pixels}$  seconds. Thus, the ADI graphics can be redrawn about once every .250 seconds. In practice, the Apple II will be much slower than this, since it would have to make computations, receive VAX input, and display all of the flight instruments, not just the ADI. Thus, a 4 Hz ADI bandwidth is quite optimistic with regard to Apple II capabilities, but it is insufficient with respect to the pilot's needs. It is our understanding that a 6 to 8 Hz bandwidth is necessary for realistic control.

While it may well be possible to write an assembly language program to make the Apple perform as desired, it would be expensive from a labor standpoint. Consequently, alternative graphics devices are being investigated. Currently, the specifications for this graphics device are that it have a high speed, parallel interface to the VAX and a display list. The parallel interface is necessary to update the display quickly. A display list is necessary so that all the dynamic elements on the screen can be changed in real-time.

It has also become apparent that seven more terminal ports will be necessary to drive the simulator displays. Three of these ports will be for output, three for input, and at least one will be needed by the experimenter to control the simulation. We would prefer to add more ports rather than take away from those we already have. Fortunately, ports cost only \$200 each in groups of sixteen.

## HUMAN PROBLEM SOLVING

As noted in the Introduction, research in this area is focusing on: 1) identification of rule-based models and, 2) alternative approaches to training and aiding. Both of these efforts are oriented toward human problem solving in dynamic environments including aviation and process control.

Mike Lewis, in conjunction with his Ph.D. thesis, is pursuing identification of rule-based models. The goal is to lessen the usual substantial subjectivity in the formulation of rule-based models by developing and testing algorithmic approaches. In this report, Mike discusses his use of such approaches for analysis of the CDTI data of Ev Palmer and his colleagues.

During the past three years, a portion of the efforts in problem solving have been focused on process control. Using a simulation called PLANT, Nancy Morris investigated the effects of types of knowledge on human performance. Annette Knaeper developed a rule-based model of human problem solving in PLANT. In comparing the behavior of this model with that of humans, Annette found that humans often did not follow their instructions, namely, PLANT operating procedures. This led to the idea of using the rule-based model for online aiding and training of operators. Annette and Nancy performed an initial empirical evaluation of this concept, the results of which are discussed in this report.

# IDENTIFICATION OF RULE-BASED MODELS

Charles M. Lewis

## INTRODUCTION

Researchers investigating pilots using cockpit displays of traffic information (CDTI) (Palmer et al., 1980 and Smith et al., 1982) have found choice of control action to depend upon individual differences as well as encounter or display characteristics. In the development of the CDTI it is important that both generalized strategies shared by all pilots and idiosyncratic choices of the few be understood. As the CDTI will likely be used in conjunction with a collision avoidance system (CAS) it is important to gauge the influence of the display on pilot maneuvers so that advisories can be formed which are both consonant with pilot strategy and avoid conflicts among strategies. The radar assisted collisions discussed by Curry (1972) demonstrate the danger of introducing such technology without consideration of operator strategies.

While established policy capturing methods exist for examining the influence of variables on decisions, they fail to elucidate what the influenced policy actually was. The present work employs pattern generalization techniques to identify a production system of rules capturing the consistencies in observed behavior. A production rule consists of two parts, a set of conditions and an action. If the conditions are satisfied the rule's action is performed. If conditions are expressed using propositional logic, descriptions are restricted to attributes (global properties of an object). Measurements commonly used in science such as height, weight, or velocity are of this type. In such cases the conditional part of the rule defines a region of the attribute space within which the rule is true (responses are of the type specified in the action part of the rule). This representation proves convenient for visualizing set theoretic relations among rules.

While this formulation of pattern generalization is similar in approach to discriminant analysis there are some important distinctions:

1. A rule describes an enclosed region of the event space rather than a partition dividing the space into two parts.
2. More than one rule may be needed to describe a response if regions in which it occurs are separated.
3. Rules may vary both in generality (size of region) and selectivity (accuracy of discrimination).

This report describes an application of pattern generalization to identification of pilot strategies.

#### CDTI DATA

Data from an experiment by Palmer (1983) investigating the effects of information quality and intruder characteristics in the use of a cockpit display of traffic information instrument has been reanalyzed using pattern generalization techniques.

In this experiment sixteen pilots "flew" sixteen programmed encounters under three display conditions. Pilots were instructed to maintain a steady course, using the autopilot unless they received a threat advisory. In response to the threat they were to maneuver to maintain a horizontal separation of greater than 1.5 nm and a vertical separation in excess of 500 ft. They were advised that an appropriate strategy was to maneuver so that the intruder would pass further away but in the same orientation at the point of closest approach.

In the least informative condition the display portrayed the relative positions of the ownship and the intruder along with tags showing their altitudes. The predictive display provided ground referenced predictors showing predicted positions of the ownship and intruder as well as a tag showing the intruders projected altitude at time of closest approach. In the third condition noise was introduced into the predictive display

#### EXPLANATORY VARIABLES

In this analysis encounter variables, describing the physical relationship between the intruder and ownship which the pilot is instructed to control, were differentiated from experimental variables. Five encounter variables were used. Four describe the relative positions of the aircraft at their point of closest approach as projected at time of alarm. The fifth measure, intruder vertical velocity, remains constant throughout

the encounter.

hpass-horizontal passing position= behind, intercept, or in front.

hsep-projected horizontal separation= very near(0-.24nm), near(.24-1nm),  
or far( > 1nm).

vcross-vertically crossing trajectories= no,yes.

vsep-projected vertical separation= very near(0-140'), near(140-350'),  
or far( > 350').

vveloc-intruder vertical velocity= zero or non-zero.

Of the sixteen programmed encounters, encounters 7 and 8 which introduce crossing angle between the aircraft as a variable were excluded. Encounters 11-16 which involve abrupt changes in intruder course or introduction of intruder in near proximity to ownship, invalidating projections made at time of alarm, remain unanalyzed as in Palmer's (1983) report.

Two non-encounter variables were considered, display type and pilot. Display type in conjunction with the encounter variables describes the stimuli under which a decision is made. Inclusion of pilot identification in the generalization introduces individual differences.

#### RESPONSE VARIABLE

Pilots' responses were represented in terms of maneuvers toward or away from the intruder along a dominant axis. The dominant axis was determined by comparing the horizontal and vertical magnitudes of a maneuver to the respective tolerances which the pilots had been instructed to maintain. Five response classes result: no action, vertical-toward, vertical-away, horizontal-toward, and horizontal-away.

#### PERFORMANCE MEASURES

Nonparametric measures of association tau-b, the ratio of between groups sum of squares to total sum of squares, and PRE, the reduction in error relative to assigning the modal response to all cases, provide measures of rule performance which consider both coverage and discrimination. Tau-b provides a nonparametric analog to a squared correlation with values under .1 indicating a relatively weak association

(corresponding to  $r < .30$ ) and those over  $.5$  a relatively strong one (corresponding to  $r > .70$ ). Using tau, single rules are evaluated by comparing the distribution of response classes within the rule with that of the remainder of the cases. This provides a measure (barring intersections) of described variance contributed by that rule to its rule set. Rule set performance may be evaluated relative to the situations in which it applies or to the entire range of examples. When restricted to applicable regions, tau may be interpreted as a measure of the extent to which the rule set describes identified consistencies. When evaluated relative to the entire space, an additional "response category" formed by uncovered observations is required. In this case tau may be considered a measure of rule set performance relative to arbitrarily chosen examples.

#### Aq ALGORITHM

A pattern generalization program, INDUCE 3 (Hoff, et al. 1983), was obtained for use in rule identification. In this analysis only the VLI (Aq) subprogram which identifies rules in propositional logic was employed.

The Aq algorithm generates a set of putative rules which match a particular positive example and exclude all negative examples. The rule which matches the most additional positive examples is retained. At each iteration successfully matched examples are removed from consideration. The process terminates when all non contradictory positive examples have been matched. Although previously matched examples cannot contribute to the retention of rules, they become "blanks" in the space, which being neutral, may become part of subsequent generalizations. The resulting rules may overlap substantially. If "rectangular rules" were identified for figure 1, three rules would be found: Rule-1=(2,3,4,5), Rule-2=(2,3,4,6,7,8), Rule-3=(1,3,7). As Rule-1 substantially describes this space with little non redundant contribution from rules 2 or 3, a parsimonious description may allow the smaller regions, 1,6,7, and 8 to go undescribed. Under other circumstances collapsing across an explanatory variable to produce a more general rule making occasional errors may be the choice dictated by parsimony.

## EFFECTS OF NON-ENCOUNTER VARIABLES

The complete set of rules generalized using the Aq algorithm provides an upper bound on the consistency with which the responses can be associated with the explanatory variables employed. The contribution of a variable may be examined by comparing performances between rule sets generalized with and without that variable. While modest improvement will be obtained from an additional variable based on an increase in degrees of freedom, major improvements in description will mirror the "influence" of that variable on pilot decisions. An index to the relative contribution of an explanatory variable may be found by rank ordering rules by performance. The relative performance for same sized sets of rules can then be compared for rule sets of varying sizes.

A generalization based on encounter variables alone produced 230 correct matches with 154 errors resulting in  $\tau\text{-}b=.18$ . If individual differences among pilots are considered as well, correct matches rise to 324 with 60 errors and  $\tau\text{-}b=.61$ . Less improvement is found in the generalization based on encounter variables and display types: correct matches=254, errors=130,  $\tau\text{-}b=.27$ . If display type and individual differences are entered into the generalization simultaneously only one error occurs yielding a  $\tau\text{-}b$  of .99.

VARIABLES	No. Rules	Hits	FAs	Tau-b	PRE
ENCOUNTER	20	230	154	.18	.38
ENCOUNTER + DISPLAY	42	254	130	.27	.48
ENCOUNTER + PILOT	85	324	60	.61	.76
ENCOUNTER + PILOT + DISPLAY	114	383	1	.99	.99

Considering performance as a function of the number of rules reveals the same ordering of effects as found for the complete rule sets. The steeper slope of the generalization including individual differences and display type indicates the importance of their interaction in describing control strategy. Individual differences appear the stronger of the factors, halving the number of errors found in a generalization based on

encounter variables alone. Pilots appear to develop individualized strategies which are influenced in similar ways by the type of display being used. Individual differences in the adaptation of control strategy to display, however, appear necessary to account for pilot behavior in detail.

#### IDENTIFICATION OF STRATEGIES

Examining the effects of non-encounter variables by comparing the performance of complete rule sets relies on the Aq algorithm's capability of finding a set of rules embodying whatever consistencies are present in the data. In this usage, ability to phrase noncontradictory rules is more crucial than their generality. When used to identify strategies, however, the generality and performance of particular rules or families of rules becomes of primary importance.

General strategies tend to be somewhat broader than absolute noncontradiction requires. Particular pilots, displays, and encounters often demonstrate slight variations on more basic strategies. In extracting strategies from a rule set it is necessary to consider a number of explicit trade-offs:

1. Discrimination- The strategy should make few false matches
2. Generality- The strategy should apply to many of the examples
3. Uniqueness- Multiple identified strategies should not match the same examples
4. Coverage- Selected strategies should cover a substantial portion of the examples
5. Parsimony- Only a small number of strategies should be identified

In spatial terms these criteria call for partitioning a large part of the attribute space (coverage) into a small number (parsimony) of large (generality), homogeneous (discrimination), non-intersecting (uniqueness) regions. These goals are often conflicting. As the size of regions (and concomitantly coverage of the rule set) increases, so does the likelihood of matching negative examples or intersecting neighboring regions. Identification of strategies requires selection of a subset of representative rules which "best" meet these criteria.

While the appropriate quantification of these criteria is not apparent it is not necessary for a rough identification of major strategies. Selection of a subset of rules from major regions of homogeneity requires only that the analyst simultaneously consider rule performance and region. Once selected, the performance of the reduced rule set can be evaluated and its usefulness as an abstraction of major consistencies in the observations appraised. Other possible selections do not invalidate this choice but merely vary the fineness of detail in exchanging generality for discrimination or parsimony for coverage. The resulting rule sets provide production system models of the observed behaviors. Conditions under which consistent responding failed to occur can be identified as well.

#### RULE TREES

Since rules may be refinements of one another or otherwise share observations it is necessary to consider rule sets in a way making their redundancy explicit. This is facilitated by representing rules in trees in which successors are subsets of their predecessors. Rules below a selected rule then describe subregions of that rule while the rule, itself, demarcates a subregion of the rules above it. Rules which are not subsets of any other rule form roots.

Well developed tree structures are typical of major strategies. Roots are found in generalizations collapsed across non-encounter variables while more specific generalizations provide refinements and variations on this basic theme attributable to particular pilots, displays, and encounters. Solitary roots by contrast tend to delimit smaller, less populous regions of the attribute space.

Rules from all generalizations, with  $\tau > .01-.02$  to exclude those covering only two or three events, were assembled into rule sets for each control action. Trees were then generated for each response type. A rule was considered a subset if:

1. proper subset- its conditions were a subset of its predecessor's
2. phenotypic subset- all events covered were also covered by its predecessor

3. intersecting subset- 90% of events covered were also covered by its predecessor

While representing rules within trees clusters those most closely associated, even roots may share substantial numbers of observations in common. In selecting rules depicting general strategies it is also necessary to consider the uniqueness of these rules which are not quite so closely related. This overlap can be conveniently expressed in an intersection matrix whose entries are the number of common observations for the rules appearing in its indices. Although less complete in its depiction than the rule tree which represents relations directly in the attribute space, the intersection matrix provides a convenient means for representing more isolated regions. In identifying major strategies the analyst may use the structural information provided by rule trees along with the more complete picture of intersections supplied by the matrix to choose rules from among branches, between trees, and among roots. This task will generally prove less formidable than it sounds since a major strategy will usually spawn a tree with a good representative(s) near its root while isolated roots typically have low coverage and may be disregarded.

#### RESULTS

A set of 9 rules were selected from the generalizations based on rule trees and intersection matrices. The selected rule set covers 44% of the sampled event space with 143 correct matches and 24 errors yielding  $PRE=.77$  and  $\tau=.61$ . When performance is considered relative to the entire event space these figures become: correct matches=213 errors=171 with  $PRE=.32$  and  $\tau=.27$ .

Two rules describe conditions for taking no action, three for turning vertically away, and four for turning horizontally toward. Turning vertically toward the intruder occurred very rarely (12 out of 384 encounters) and so was not modeled. The horizontal away response accounting for 70 of the 384 encounters also was not represented. Although 73% of these occurrences are successfully described by a set of 29 horizontal-away rules with only 21 errors, these rules have uniformly small coverage and low overlap. Over half of the horizontal-away rules were restricted to groups of five or fewer pilots indicating the idiosyncratic (or coincidental) nature of this response choice. The

overall inconsistency in the choice of this response is revealed in the rule trees for horizontal-away where 27 of 29 rules stand alone if a 90% inclusion criterion is applied. To consider pilot strategy it is necessary to examine the rules, themselves, in greater detail. This will be done for each response class.

#### NO ACTION

##### Rule No. 1

[Pilot=7,8,9,10,11,12,13,14,15,16] & [Horizontal passing position=intercepting or in front]  
 & [Projected horizontal separation=far]  
 correct matches=22, errors=7, PRE=.09, tau-b=.07

##### Rule No. 2

[Pilot=3,4,7,8,11,14,15,16] & [Projected horizontal separation=far] &  
 [Vertical crossover=true]  
 correct matches=18, errors=6, PRE=.07, tau-b=.05

#### RULE SET SUMMARY FOR NO-ACTION

correct matches=35, errors=11, PRE=.14, tau-b=.11

The essential condition for eliciting no response appears to be a large projected horizontal separation. Thirty of the 36 rules found for no-response were refinements of this condition, [Projected horizontal separation=far]. Standing alone this condition produces 43 correct matches with 52 errors. The two rules selected miss 8 of these matches but result in 41 fewer errors.

Both individual differences and other aspects of the encounters appear responsible for the increase in selectivity. In the first rule, encounters in which the intruder would pass behind are excluded. This finding is consistent with (O'Connor et.al., 1980) and findings in relation to the horizontal-toward response in this study, that pilots tend to maneuver in a way to cause intruders to pass in front. Individual differences have an equally clear influence. Pilots 7,8,11,14,15, and 16 appear in both of the selected rules. If [Projected horizontal separation=far] is constrained to this group of pilots, correct matches are reduced by only 44% while errors decline by 79%. If the selected rules were restricted to these pilots, selectivity again improves with correct matches declining from 35 to 22 and errors from 11 to 5.

While pilots made no response on only 48 out of the 384 encounters, patterns are found for this choice. A relatively small group (6-8) of pilots account for almost all occasions on which a constant course was maintained. This choice was made appropriately for large horizontal separations but failed to occur when the major separation was vertical or the intruder was oriented to pass behind.

#### VERTICAL AWAY

##### Rule No. 3

[Pilot=2,6,7,8,10,11,12,15] & [Projected vertical separation=near] &  
[Vertical velocity=zero]  
correct matches=43, errors=3, tau-b=.09

##### Rule No. 4

[Pilot=6,7,11,13,16] & [Projected horizontal separation=near] &  
[Projected vertical separation=very near or near] &  
[Vertical velocity=zero]  
correct matches=27, errors=1, tau-b=.06

##### Rule No. 5

[Display=no predictor] & [Vertical velocity=zero]  
correct matches=27, errors=5, tau-b=.05

#### RULE SET SUMMARY FOR VERTICAL-AWAY

correct matches=62, errors=7, tau-b=.12

Rules for the vertical away response are contained within the portion of the space in which the intruder is approaching at a constant altitude. Seventy-six of 134 encounters responded to with a vertical-away response were of this type. While 34 additional encounters are covered by 13 more rules in which this condition is not explicitly expressed, their observations fall largely within the constant altitude region. The failure to find strong rules covering the 54 encounters in which the intruder changed altitude indicates an inconsistent usage of the vertical-away response under these conditions. Rules 3 and 4 contain proximity conditions and apply to all displays. In rule 5 both proximity and individual differences are dropped. In the absence of predicted separation pilots chose a vertical away response when confronted with an intruder at constant velocity regardless of the actual threat. Examination of these

rules indicates that, for these encounters, projected proximity information influenced the decision to respond but not the response chosen. Pilots' choice of the vertical-away response appears limited to the constant altitude intruder although a strategy of increasing vertical separation would apply to vertically moving intruders as well. The presence of predicted altitudes does not appear to influence this decision. While the vertical-away response was the modal response in this study its association with a clearly discriminable form of separation information rather than projected separations provided by the predictor displays may indicate some difficulties in the use of this information to guide control actions.

#### HORIZONTAL TOWARD

##### Rule No. 6

[Pilot=7,11,12,14] & [Display=no predictor or predictor] &  
 [Projected horizontal separation=very near or near] &  
 [Vertical crossover=no] & [Vertical velocity=not zero]  
 correct matches=14, errors=0, PRE=.06, tau-b=.03

##### Rule No. 7

[Pilot=6,11,12,14] & [Projected horizontal separation=near] &  
 [Projected vertical separation=far] & [Vertical velocity=not zero]  
 correct matches=9, errors=0, PRE=.04, tau-b=.02

##### Rule No. 8

[Pilot=4,5,11,12,15] & [Projected horizontal separation=very near] &  
 [Vertical velocity=not zero]  
 correct matches=19, errors=6, PRE=.07, tau-b=.03

##### Rule No. 9

[Pilot NE 5,6,10] & [Display=no predictor] & [Passing position=intercept or in front]  
 & [Projected horizontal separation=near] & [Vertical velocity=not zero]  
 correct matches=15, errors=0, PRE=.06, tau-b=.04

#### RULE SET SUMMARY FOR HORIZONTAL-TOWARD

correct matches=46, errors=6, PRE=.18, tau-b=.10

Rules for the horizontal-toward response are contained within the complementary "changing intruder altitude" portion of the event space. This factor rather than proximity or relative orientation appears crucial in the choice between horizontal and vertical responding. While the

vertical-away response was chosen consistently throughout the constant altitude region, the choice of the horizontal-toward response is less monolithic. As noted in the discussion of "no-response", this choice occurs almost exclusively (98%) in this region. Similarly 83% of the vertical-toward, 43% of the vertical-away, and 93% of the horizontal-away responses occur in encounters in which the intruder is changing altitude.

Rules 6, 7, and 8 are refinements based on individual differences of a strategy of turning toward intruders who are laterally close and changing altitudes:

[Projected horizontal separation=very near or near] &  
 [Vertical velocity=not zero]  
 correct matches=85, errors=107, PRE=.12, tau-b=.05

The poor performance of the rule expressed in this general way indicates this strategy is followed by only a small group of pilots. The improved selectivity of rules 6 and 7 is attributable primarily to pilots 11, 12, and 14. The general rule restricted to these pilots:

[Pilot=11,12,14] & [Projected horizontal separation=very near or near]  
 & [Vertical velocity=not zero]  
 correct matches=28, errors=8, PRE=.08, tau-b=.04

accounts for 75% of the encounters covered by rules 6 and 7 and represents a major improvement in selectivity over its unrestricted form. Rule 8, another refinement of the general strategy which restricts the rule to intruders at the closest proximity, is followed by a larger group of pilots. None of these rules shares as many as 60% of its observations with the rule embodying the recommended strategy for a horizontal-toward response:

1. There is a threat
2. Maneuvering horizontally toward the intruder will maintain the aircrafts' relative horizontal positions and increase horizontal separation at point of closest approach.

[Passing position=intercept or in front] &  
 & [Projected horizontal separation=very near or near]  
 correct matches=65, errors=79, PRE=.02, tau-b=.03

Rule 9, by contrast, is a refinement of the recommended strategy fitting most pilots using the display without a predictor.

## DISCUSSION

Within the range of encounters examined, the vertical movement of the intruder appears the most crucial factor in determining the pilot's dominant response. Under conditions in which the intruder approached at a constant altitude pilots under all displays, with few individual differences, and with little regard to the degree of threat, maneuvered vertically away. This strategy follows the principle of least effort in limiting the decision to a single dimension (vertical velocity) and producing a response which increases separation at point of closest approach under all conditions. While ensuring success at the pilots' primary task of avoidance, this strategy may run counter to the secondary task of maintaining course in the face of nonthreatening encounters. This shortcoming is highlighted by noting that of 48 occasions on which the pilot did not maneuver only one occurred under these conditions.

When the intruder was changing altitude the vertical response dimension was largely ignored accounting for the dominant response on only 24% of such occasions. As in previous studies (Palmer et al. 1981, Ellis and Palmer 1982, Smith et al. 1982,1984) horizontal-toward were preferred to horizontal-away responses. Palmer et al. 1981 have attributed this tendency to the pilots' desire to maintain visual contact with the intruder while Ellis and Palmer (1982) have suggested they desire, instead, to minimize the time to resolution of the conflict by passing behind the intruder. Regardless of the motivation, this effect is found consistently in CDTI studies and should be considered in assessing the usefulness of such displays. Smith et al. shed additional light on this preference, finding that encounters rated as less threatening showed a stronger turning-toward tendency. Rules identified for the horizontal-toward response support this view showing a general preference for the horizontal-toward response while using predictor displays which allowed a clear view of conflict resolution but limiting the response to the more conservative recommended strategy when the display lacked predictors.

As found in earlier studies (Smith et al., Palmer et al., Ellis and Palmer) large individual differences were noted among pilot's strategies. The most nearly universal decision was the choice of the vertical-away maneuver under conditions in which it unambiguously increased separation.

The rules identified suggest that vertical information may not be presented in the most useful manner. None of the nine selected rules contain any reference to this relation although it contributes as much to achieved separation and collision avoidance as the horizontal dimension. This neglect is further reflected in the pilots' overall preference for horizontal maneuvers. Smith et al. have suggested the preference for horizontal responses may be due to FAA regulations, comfort, safety or fuel conservation but the absence of vertical information from decision rules suggest the bias may more likely be due to the superior display of horizontal traffic information.

The finding that pilots using predictive CDTI displays were more likely to proceed with conflict resolution by turning toward the intruder than following the recommended strategy reinforces concerns aired by Palmer et al. (1981), Lester and Quan (1983) and others that CDTI in some instances may actually make collisions more likely. Pilots, themselves, are not immune to this fear. The October 28, 1984 New York Times observes that, "The Airline Pilots Association has been especially insistent that the devices must ultimately be able to recommend a horizontal right turn or left turn maneuver in addition to a vertical maneuver." Earlier analysis of this data (Palmer 1983) indicates that the noiseless predictor display led to fewer positive CAS advisories and smaller maneuver magnitudes while the predictorless display resulted in smaller achieved separations and less frequent agreement with the recommended strategy. The present investigation suggests that the superiority in performance on the predictor displays results from improvements in execution rather than fundamental shifts in strategy. For one group of pilots, in fact, consistent violation of the recommended strategy was linked with the use of the noiseless predictor display. While the most widely employed strategy observed was the vertical away response to a constant altitude intruder, vertical responses were generally avoided under other conditions. Since projected altitudes at closest point of approach provide information unavailable from rapidly updating data tags, the failure to find a related consistency in pilots' responses suggests some difficulty in abstracting or using this more detailed altitude information as it is presented.

## REFERENCES

1. Curry, R.E. Will the ATSD precipitate display-assisted collisions? in H.G. Weiss and R.W. Bush (eds.), The Role of an Airborne Traffic Situation Display in an Evolving ATC Environment (PB-215-714). Cambridge, Ma: MIT, 1972.
2. Ellis, S.R. and Palmer, E.A. Threat perception while viewing single conflicts on a cockpit display of traffic information (Report No. 81341). Moffett Field, Ca.: NASA-Ames Research Center, 1982.
3. Hoff, William A., Michalski, Ryszard S., Stepp, Robert E. "INDUCE 3: a program for learning structural descriptions from examples", in draft form, Department of Computer Science, University of Illinois at Urbana-Champaign, 1983.
4. Lester, P.T. and Quan, E.E. "The cockpit display of traffic information and the threat alert and collision avoidance system integration: a review", Proceedings of the Second Symposium on Aviation Psychology, Columbus, Ohio: Aviation Psychology Laboratory, Ohio State University, 1983.
5. O'Connor, S., Jago, S., Baty, D. and Palmer, E. The effect of viewing time, time to encounter, and practice on perception of aircraft separation on a cockpit display of traffic information (Report No. 81173). Moffett Field, Ca.: NASA-Ames Research Center, 1980.
6. Palmer, Everett "Conflict resolution maneuvers during near miss encounters with cockpit traffic displays", Proceedings of the Human Factors Society-27th Annual Meeting, Santa Monica, Ca.: Human Factors Society, 757-761, 1983.
7. Palmer, E.A., Jago, S.J., Baty, D.L., and O'Conner, S. Perception of horizontal aircraft separation. on a cockpit display of traffic information. Human Factors, 605-620, 22, 1980.
8. Palmer, E., Jago, S., and DuBord, M. "Horizontal conflict maneuvers with a cockpit display of traffic information", Proceedings of the Seventeenth Annual Conference on Manual Control, Los Angeles, Ca.: UCLA, 1981.

9. Smith, J.D., Ellis, S.R., and Lee, E. Avoidance maneuvers selected while viewing cockpit traffic displays (NASA TM-84269). Moffett Field, Ca. 1982.
10. Smith, J.D., Ellis, S.R., and Lee, C.E. "Perceived threat and avoidance maneuvers in response to cockpit traffic displays", Human Factors ,33-48, 26, 1984.

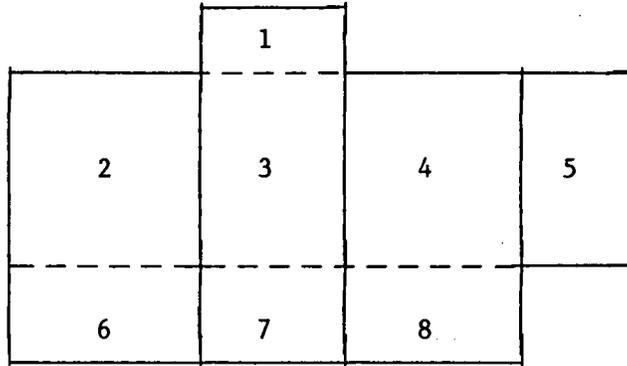


Figure 1

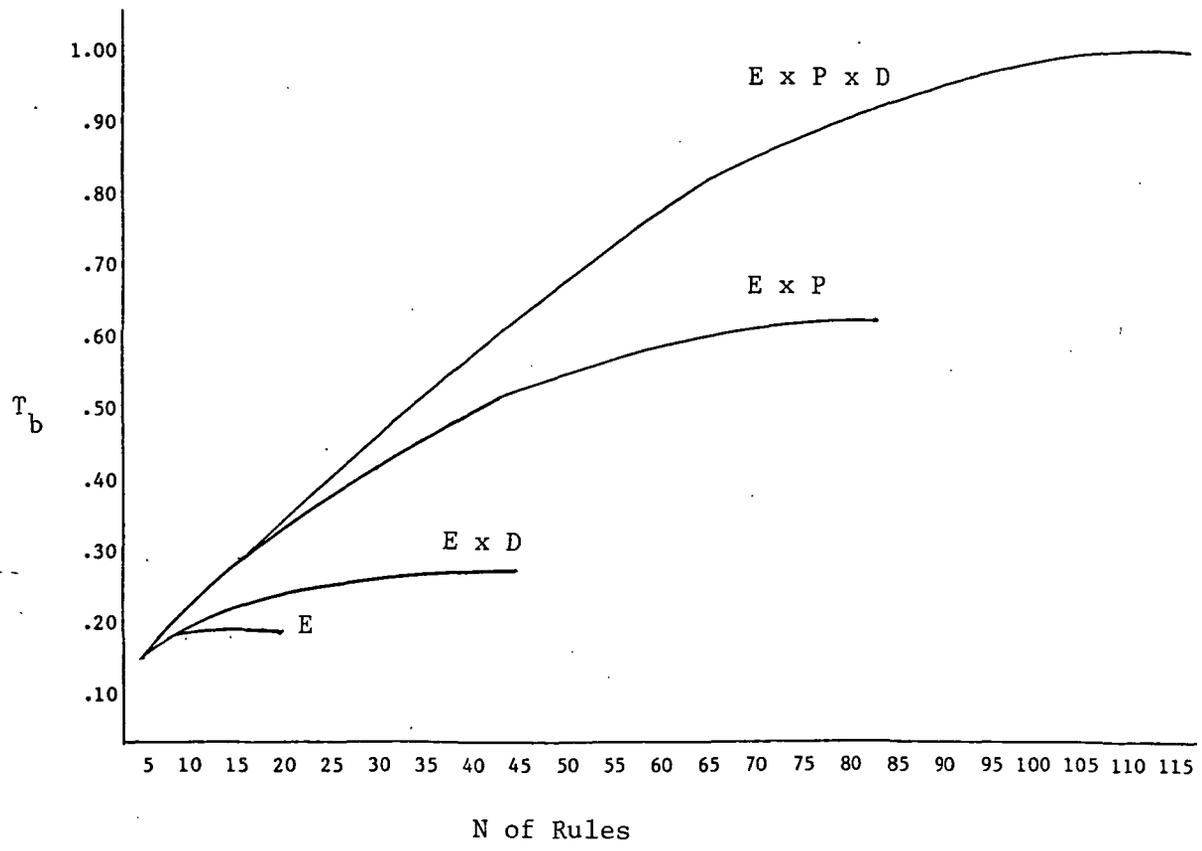


Figure 2

E - encounter  
 D - display  
 P - pilot

A MODEL-BASED APPROACH FOR ONLINE AIDING  
AND TRAINING IN PROCESS CONTROL

Annette Knaeuper  
Center for Man-Machine Systems Research  
Georgia Institute of Technology  
Atlanta, Georgia 30332

and

Nancy M. Morris  
Search Technology, Inc.  
Norcross, Georgia 30092

ABSTRACT

This research addressed the feasibility of adapting an existing rule-based system as an online "coach" for controlling PLANT, a simulation of a generic process plant. KARL, a rule-based model capable of controlling PLANT, was adapted to provide three types of information to subjects: 1) situation assessment (i.e., which operational procedure, if any, was applicable for a given situation); 2) guidance in following procedures (i.e., feedback whenever subjects' actions were inconsistent with available procedures); 3) performance feedback (based upon changes in the system's stability). Subjects received this information online while controlling PLANT. Compared to subjects in an earlier experiment who controlled PLANT without the benefit of the coach, these subjects maintained a generally more stable system, scored higher on a paper-and-pencil test of system knowledge, and were more successful in diagnosing an unfamiliar failure of the PLANT safety system. Careful analysis of these results in light of previous research with PLANT indicated that the reasons for these differences were not as straightforward as they might appear. This experiment is viewed as illustrating potential benefits and subtleties of using a rule-based model as an online coach.

## INTRODUCTION

As systems increase in complexity, the question of how persons should be trained to operate them becomes more important. The amount of training required for someone to become proficient at controlling a complex system may be quite extensive, and it is necessary to consider a number of issues when developing such a training program. These issues include the content and format of instructional material and the structure of the program. Because of inherent human limitations, it may also be necessary to consider provision of some kind of performance aid, in addition to appropriate training.

Many reports are available which directly or indirectly address issues relevant to training (Morris & Rouse, 1984b). Some are directed at obtaining an understanding of how people solve problems, either in the laboratory or in contact with an actual system. Others investigate the effects of various training approaches upon performance. Often there is a discussion of the human's "mental model" of the system being controlled (Rouse & Morris, 1984).

One study in particular served as a basis for the present research. Morris investigated the effects of different types of instruction upon subjects' ability to control PLANT, a computer-based simulation of a generic fluid production process

(Morris, 1983; Morris & Rouse, 1984a). The PLANT operator's task is to supervise the flow of fluid through a series of tanks interconnected by valves so as to maximize production. This may be done by opening and/or closing valves and adjusting input and output, via commands entered at the terminal keyboard. A number of failures may occur in PLANT, so there are several diagnostic and repair commands available as well.

The primary comparison in Morris' research was between two different types of instruction: 1) operational procedures, and 2) a description of dynamic principles and functional relationships in PLANT. Four groups of subjects were compared, distinguished on the basis of the combination of written instructional materials they received (i.e., principles, procedures, neither principles nor procedures, or both principles and procedures). Instruction was found to have no effect upon subjects' achievement of the overall goal of production, in that there were no differences between groups with respect to this measure. However, those groups receiving procedures were found to control the PLANT in a more stable manner, even though all groups had been told to maintain stability.

An interesting aspect of this research was an investigation of subjects' ability to deal with two unfamiliar failures: a tank rupture, and failure of the PLANT safety system. (The failures were unfamiliar in that, although subjects knew they

could occur, they had not experienced them before.) Almost all subjects repaired the tank rupture; however, only half of the subjects in each group successfully diagnosed the safety system failure. This was surprising, because subjects with an understanding of the functioning of the system (as described in the principles) should have been better able to make that diagnosis.

As a result of these findings, it was suggested that one of the reasons a knowledge of principles failed to help many subjects deal with the unfamiliar failure was that those people did not realize that they were in an unusual situation, and thus did not realize that they should use their knowledge. In other words, they failed to make an accurate assessment of the situation. This notion was indirectly supported by the fact that those persons who did repair the unfamiliar safety system failure also maintained a more stable system in general; since the effect of the safety system failure was to make the PLANT appear more unstable, maintaining a stable system may have enabled subjects to detect the presence of an unusual situation more readily.

Some useful insights into subjects' behavior were gained by comparing their performance to that of KARL (Knowledgeable Application of Rule-based Logic), a model capable of controlling PLANT (Knaeuper, 1983; Knaeuper & Rouse, 1984). KARL is a

rule-based model patterned after a general model of human problem solving proposed by Rouse (1983), which suggests that problem solving is accomplished in three stages: 1) recognition and classification, 2) planning, and 3) execution and monitoring. These three stages essentially define KARL's structure. When controlling PLANT, KARL accesses a knowledge base consisting basically of information contained in written information available to subjects (i.e., operational heuristics and procedures, and information about dynamic principles and functional relationships).

When the performance of subjects and KARL was compared, it was noted that KARL consistently achieved higher production and maintained a more stable system than did subjects. It was also interesting to examine differences in the courses of action chosen by subjects and KARL in solving problems in PLANT. Basically, two rather systematic differences were found. First, the levels of system input and output chosen by subjects were not as high as those chosen by KARL (and suggested by procedures); subjects were more conservative in this respect. Second, KARL adjusted input and output much more frequently than did subjects; this reflected heuristics within KARL which were directed at maximizing production, which were not a part of operational procedures.

Considering some of the apparent difficulties experienced by

subjects in making an accurate situation assessment and following procedures, and the benefits derived from using KARL as an off-line analysis tool, an idea emerged. Why not make it possible for KARL to analyze subjects' actions online and provide advice, thus functioning as an online "coach"? It seemed that such an approach could prove to be useful for both training and aiding.\*

#### DESCRIPTION OF THE COACH

In light of the factors noted above, the decision was made to provide subjects with three types of information. In the context of PLANT, this information was displayed on the terminal near the area where normal operating messages were displayed. The first type of information was related to situation assessment. Specifically, a message informing the subject which procedure was currently applicable was shown (e.g., "Procedure 5"). If no procedure applied, the following message was displayed: "No procedure applicable; Normal tuning".

Subjects also received guidance in following procedures. KARL monitored subjects' actions, and provided feedback if a given action was inconsistent with the applicable procedure. For

---

\* Of course, one could view this approach as simply a special case of "expert systems". This issue is discussed later in this paper.

example, the following message might appear: "Your action (cva,e)\* is inconsistent with Procedure 5. Keep all valves open until the system is stable again. Type 'y' for change." As may be ascertained from the last portion of the message, subjects had the option of overriding KARL or changing their actions to be consistent with KARL's recommendations.

The third type of information supplied by KARL was performance feedback, or information about the degree to which subjects' actions were succeeding in remedying problems in the system. This information was supplied because of the length of time required for the consequences of actions to become manifest. These messages were based upon changes in PLANT stability over a period of 10 time units, and consisted of the following: "Instability extreme", "Instability excessive", or "Instability improving".

The process of enabling KARL to supply such messages was relatively straightforward. However, when an attempt was made to control PLANT with KARL as an assistant, a number of problems became apparent. For example, KARL's advice as to what actions should be taken was not always consistent with procedures. This could be attributed to the nature of KARL's approach to PLANT.

---

\* cva,e = close the valve between tanks a and e

Although the information in the procedures was contained in KARL's knowledge base, KARL also employed several heuristics when controlling PLANT, which occasionally preempted the action recommended in procedures.

Another problem was related to KARL's situation assessment. During the course of PLANT operation, situations would occasionally arise which were "borderline" conditions with respect to the applicability of various procedures. KARL's decisions as to which procedure applied were based upon fixed values of state variables. In borderline situations, normal fluctuations of these state variables caused KARL to change the situation assessment message rather frequently (e.g., every other time unit).

A third source of difficulty was KARL's "persistence" in reporting actions which were inconsistent with procedures. The PLANT operator was given the option of overriding KARL and implementing an action against KARL's recommendations. However, the consequence of thus failing to conform was to receive another message. KARL did not know how to concede; in short, KARL was a nag.

These problems were remedied in two general ways. First, it was necessary to inhibit the display of all messages which were not procedure-oriented. Second, thresholds for prompts were

incorporated. For example, if a subject failed to comply with one of KARL's suggestions, KARL did not make the same suggestion again for five time units. As another example, "hysteresis" was introduced into the situation assessment thresholds to avoid the aforementioned problem of borderline conditions.

An experiment was conducted to evaluate the effectiveness of KARL as an assistant. Two general issues were of interest: 1) the feasibility of adapting a rule-based system (which was not originally designed as an aid) to support human problem solving, and 2) the effects of an online coach upon humans' performance.

## METHOD

### Subjects

Junior and senior undergraduates at Georgia Institute of Technology served as paid volunteer subjects. All eight of them were majors in industrial and systems engineering, and had completed courses in physics, dynamics, calculus, and differential equations.

### Experimental Procedure

The experimental procedure in this experiment was almost identical to that used in the research described earlier (Morris, 1983; Morris & Rouse, 1984a). Training provided to subjects in this experiment was equal to the group receiving instruction in

both principles and procedures in the earlier experiment, with the exception that aiding was available.

Subjects served in a total of 13 sessions each, with the average length of each session being approximately 60 to 75 minutes. Generally, training was accomplished during the first eight sessions, in which subjects read instructional materials and practiced controlling PLANT. A discussion of principles governing PLANT was provided during session 3, and operational procedures were made available for the first time in session 5. KARL was used as an online coach during sessions 5-8, and supplied the three types of aiding information described earlier.

Sessions 9-13 were considered experimental sessions, in that no further instruction was provided by the experimenter, and no questions from subjects were answered. As with the earlier experiment, unfamiliar situations (i.e., a tank rupture and a safety system failure) were introduced in sessions 10 and 12, which were counterbalanced across subjects. The coach did not provide guidance in following procedures during sessions 9-12; subjects received only information related to situation assessment and overall performance feedback. No information from the coach was available during session 13. At the end of session 13, subjects completed a paper-and-pencil test of knowledge about PLANT and the coach, based upon material contained in the written instructions.

## RESULTS

In order to assess the effects of aiding, the performance of subjects in this experiment was compared via analysis of variance to performance of the group receiving both principles and procedures in the earlier PLANT research. (In the following presentation, these groups are referred to as the aided and unaided group, respectively.) Thus, performance measures were used as dependent variables in two-way analyses with one between-subjects factor (aiding) and one within-subjects factor (session).

As with the earlier research, the experimental manipulation had no significant effect upon total production achieved, although the mean for the aided group was slightly higher (344.6 vs. 320.2 units of production per time unit). There was also no significant effect of aiding on the number of automatic valve trips experienced (an indication of PLANT stability). However, as with total production, the mean for the aided group was slightly better (i.e., lower) (0.497 vs. 0.605 trips per time unit).

Aiding also failed to have a significant effect upon another measure of PLANT stability: variance of fluid levels in the system. Once again, the trend was in the expected direction, in that the mean for the aided group was lower (12.44 vs. 15.27).

Two performance measures were significantly affected by aiding. Aided subjects kept a higher percentage of valves open (92% vs. 87%,  $p < .04$ ), and generally maintained a higher level of input into the system (116.8 vs. 106.9 units per time unit,  $p < .04$ ). The practical significance of these results is presented later.

Assessing subjects' performance during unfamiliar situations, there was no effect of aiding upon subjects' repair of the tank rupture (15 of the 16 subjects did so). However, it was found that seven out of eight subjects in the aided group repaired the unfamiliar failure of the PLANT safety system, whereas only three of the eight unaided subjects found that failure. This difference in proportions was found to be statistically significant ( $p < .04$ ).

Differences in scores on the test of PLANT knowledge were examined. Although overall scores did not differ significantly, it was found that the aided group scored significantly higher on the section of the test related to dynamic principles (83% vs. 69%,  $p < .05$ ).

Finally, the actions selected by subjects were compared to actions which would have been selected by KARL in the same situation. This comparison was similar to that reported for the earlier experiment (Knaeuper, 1983; Knaeuper & Rouse, 1984).

There was no significant difference in the degree to which actions chosen by aided and unaided subjects agreed with those selected by KARL.

#### DISCUSSION AND CONCLUSIONS

As noted in the introduction, this research was prompted by two issues: 1) the feasibility of adapting a rule-based model as an online coach, and 2) the effects of such assistance upon subjects' ability to control PLANT. With regard to the second issue, none of the statistically significant effects the coach had upon subjects' performance were related to primary performance measures. Although mean performance for the aided group was better with all measures, the only significant effects of aiding were upon the secondary performance measures of number of open valves and level of system input. These measures indicate that subjects did what they were told to do. Although all subjects (in this research and in the earlier experiment) were instructed to keep all valves open and maintain a relatively high level of input and output, apparently the coach's presence caused them to follow these instructions more closely.

Whereas it is fairly easy to provide an explanation for subjects' following instructions more closely, explaining why more subjects in the aided group were able to diagnose the safety system failure is not as straightforward. Three possibilities are suggested by the data. First, since failure of the safety

system resulted in automatic closing of valves at random, the ability to maintain more valves open in general may have assisted subjects in detecting the presence of an unusual situation. Once detected, it should have been easy to determine that the cause of the unusual situation was failure of the safety system, since only two unusual failures were possible.

Judging from the available evidence, however, it is difficult to imagine that this is a sufficient account of what happened. A look at the performance of all subjects supplied with procedures in the earlier experiment conducted by Morris (i.e., those with procedures only, and those with both procedures and principles) reveals that there was no difference in the number of valves kept open by persons who repaired the safety system and those who did not (89% vs. 88%). Additionally, a subsequent examination of logs kept by the unaided group during the time the safety system had failed indicated that at least six of the eight people felt that something was wrong; yet, only three of these successfully diagnosed the failure, and the others attributed the problem to deficiencies in their control actions.

Another possible explanation may be found in the fact that the aided group scored significantly higher on the test of information related to dynamic principles. Perhaps an increased knowledge of the functioning of the system enabled the aided group to diagnose the unfamiliar failure. This explanation also

seems inadequate. There was no difference in the test scores of unaided subjects who repaired the safety system and those who did not (69.3% vs. 69.2%).

The third explanation for aided subjects' success in diagnosing the failure of the safety system is that somehow providing them with the coach made the difference. During the session in which the safety system failed, two types of aiding messages were provided: situation assessment and performance feedback. The situation assessment consisted of informing subjects which procedure, if any, applied. There were no messages such as "unfamiliar situation". Performance feedback was related to changes in the stability of the system. When the safety system failed, it is possible that subjects received conflicting messages, such as "No procedure applicable" and "Instability extreme". Apparent conflict such as this may have served as a cue that something was wrong, and could have suggested to subjects that the problems in the system were not the result of poor control actions.

These ideas about the role of the coach in the unfamiliar situation are purely conjecture at this point. It seems likely that a combination of all of these factors (i.e., increased system stability, knowledge of the functioning of the system, and assistance in situation assessment) contributed to subjects' success. An understanding of factors affecting the human's

ability to deal with an unfamiliar event could have important theoretical and practical implications, and further investigation of this issue is warranted.

Finally, another question arises with regard to the results of this research: Why did aided subjects score higher on the test of dynamic principles? Since the primary difference in the way the two groups were treated was the presence or absence of the online coach, it would appear that this was the reason for the difference in the test scores. This is counterintuitive, however, because the focus of the aiding was on following procedures, and not on understanding the functioning of the system. Therefore, interpretation of this result must be delayed until the research can be replicated, using a larger number of subjects and controlling for potential differences in abilities.

Considering the feasibility of adapting a rule-based model as an online coach, this research has served to emphasize the complexities and subtleties of model-based online aiding and training. As noted by other researchers (Clancey & Lestinger, 1982; Jackson & Lefrere, 1984), answering the questions of what advice and feedback to provide, as well as when they should be provided, is far from straightforward. This point is particularly supported by the results reported here where subjects benefited along several dimensions by having an online coach, but did not become more like the coach in the process

(i.e., there was no increase in agreement between the subjects' and model's choices of actions). Thus, the results of being coached can be more than, or at least other than, simply gaining the coach's expertise. This has profound implications for the current view of "expert systems" as a panacea for training and aiding.

#### ACKNOWLEDGEMENTS

This research was partially supported by the Office of Naval Research under Work Unit 154-491 (Contract N00014-82-K-0487), and partially supported by the National Aeronautics and Space Administration under Ames Grant NAG 2-123.

#### REFERENCES

- Clansey, W. J., & Letsinger, R. NEOMYCIN: Reconfiguring a rule-based expert system for application to teaching (Tech. Rept. STAN-CS-82-908). Palo Alto, CA: Stanford University, 1982.
- Jackson, P., & Lefrere, P. On the application of rule-based techniques to the design of advice-giving systems. International Journal of Man-Machine Studies, 1984, 20, 63-86.
- Knauerper, A. A rule-based model of human problem solving behavior in dynamic environments (Tech. Rept. 83-3). Atlanta, GA: Center for Man-Machine Systems Research, Georgia Institute of Technology, August 1983.
- Knauerper, A., & Rouse, W. B. A rule-based model of human problem-solving behavior in dynamic environments. IEEE Transactions on Systems, Man, and Cybernetics, 1984, in press.
- Morris, N. M. Human problem solving in process control (Tech. Rept. 83-2). Atlanta, GA: Center for Man-Machine Systems Research, Georgia Institute of Technology, August 1983.

Morris, N. M., & Rouse, W. B. The effects of type of knowledge upon human problem solving in a process control task. IEEE Transactions on Systems, Man, and Cybernetics, 1984, in press. (a)

Morris, N. M., & Rouse, W. B. Review and evaluation of empirical research in troubleshooting. Norcross, GA: Search Technology, August 1984. (b)

Rouse, W. B., & Morris, N. M. On looking into the black box: Prospects and limits in the search for mental models. Norcross, GA: Search Technology, July 1984.