N77-12695

NASA Grant NSG-2119*

PILOT INTERACTION WITH AUTOMATED AIRBORNE DECISION MAKING SYSTEMS

Semiannual Progress Report

May 1976 - October 1976

William B. Rouse, Principal Investigator
Yee-Yeen Chu
Joel S. Greenstein
Rex S. Walden

Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801

*The NASA Technical Officer for this grant is Thomas E. Wempe, NASA Ames Research Center, Moffett Field, CA 94035.

CONTENTS

INTRODUCTION

This report summarizes progress of a research program whose purpose is investigation of appropriate modes of interaction between a human pilot and automated on-board decision making systems. Most of our efforts thus far have been concentrated on determining how to allocate decision making responsibility between human and computer. However, we have also initiated research on the topic of pilot problem solving in automated and semi-automated flight management systems.

In the next section of this report, Rex Walden discusses the experimental situation he has developed for study of allocation of responsibility between human and computer. He also discusses experiments in progress that are aimed at determining various pilot performance parameters with varying degrees of automation.

In the section following Walden's, Yee-Yeen Chu considers optimal allocation of responsibility between human and computer. He discusses various approaches and presents some theoretical results found in the literature. These results offer valuable insights into the allocation of decision making responsibility in flight management.

In the following section, Joel Greenstein discusses efforts aimed at developing a model of human decision making in a multi-task situation that, abstractly at least, is very similar to that being studied by Walden and Chu. The models developed with this work should be directly applicable to the design of computer aids for allocation of decision making responsibility between human and computer.

In the final section of this report, the pilot as a problem solver is discussed. The design of displays, controls, procedures, and computer aids for problem solving tasks in automated and semi-automated systems is considered.

PILOT DECISION-MAKING IN A FLIGHT MANAGEMENT SITUATION WITH VARIOUS DEGREES OF AUTOMATION*

I. Introduction

An experimental situation for study of pilot interaction with an automated airborne decision-making system has been developed. A digital computer simulation of an aircraft uses a CRT graphics system to present a cockpit-like display to an experimental subject "pilot". (See Figure 1.) The display includes standard aircraft instruments: artificial horizon, altimeter, heading and airspeed indicators, and a clock. Also displayed is a map which indicates the course the airplane is to follow. An airplane-shaped symbol indicates the airplane's actual position. A small circle moves along the mapped course and indicates the position the aircraft should have for it to be on course and schedule. Near the lower edge of the display, several dials are shown, which represent indicators for such things as fuel, electrical, or hydraulic subsystem status. The dial pointers move with dynamics provided by passing a zero-mean, Gaussian white noise input through a second-order filter.

II. The task

The pilot controls the aircraft pitch and roll dynamics with a joystick. Another control stick controls the airspeed. The pilot's task is to fly the airplane along the mapped route, maintaining a fixed altitude

---

* This section is based on the M.S. thesis proposal of R. S. Walden.

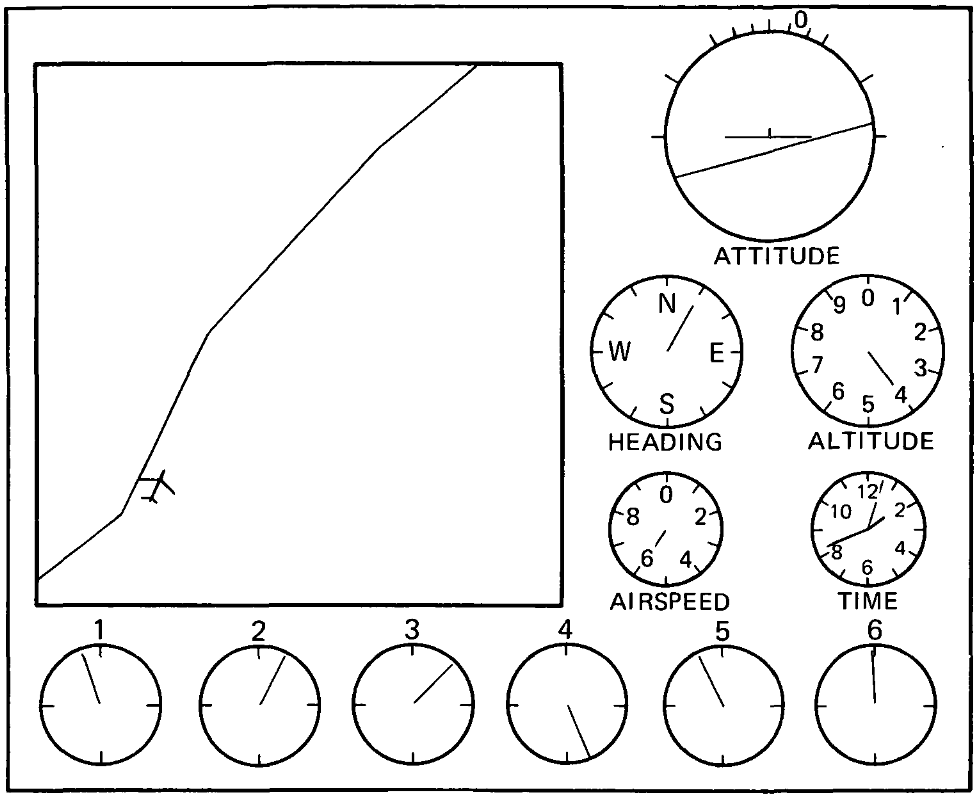

Figure 1: The Flight Management Situation

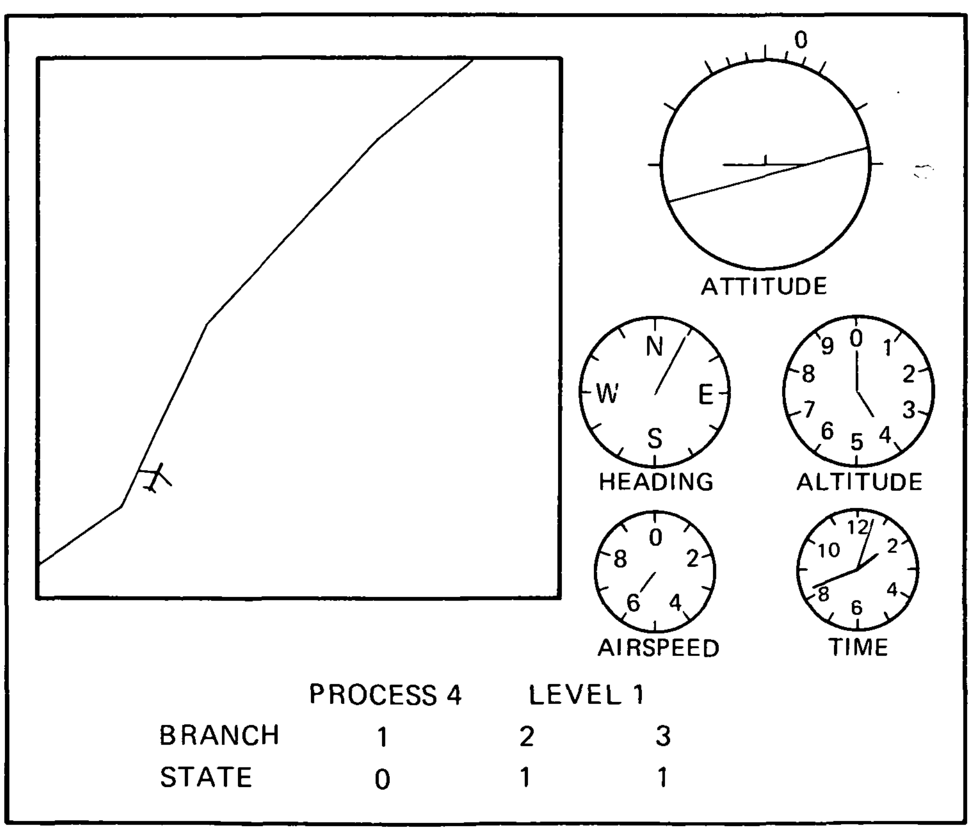

Figure 2: Display When Pilot Has Reacted to a Possible Malfunction

and stable pitch and roll attitudes. An autopilot feature is provided which causes the airplane position to coincide with the on-course indicator. While the autopilot is on, the aircraft is maintained in level flight, on course. Switching on the autopilot has the effect of resetting the airplane to the nominal, desired attitude and course, no matter what the current state of the airplane happens to be.

In addition to flying the airplane, the pilot must monitor the subsystem indicators for possible subsystem malfunction (indicated when the needle points downward - see Figure 1). When the pilot thinks a subsystem has malfunctioned, he enters the corresponding number on a 4 x 3 numerical entry keyboard. The display shown in Figure 2 then appears.

On this new display, the subsystem dials have been replaced with two rows of numbers, one labelled 'BRANCH', the other labelled 'STATE'. This corresponds to the first level of a checklist associated with the subsystem which was selected. The pilot looks for a branch with a state of '0' and enters the branch number on the keyboard. If the checklist for that branch continues, the next checklist level is displayed. When the end of the checklist is reached, the subsystem indicators reappear, with the malfunctioning system corrected. These actions simulate the checklist procedure a pilot might perform when correcting malfunctions in a real aircraft.

III. Program description

Simple aircraft dynamics are used. The pitch rate, roll rate, and airspeed, relative to aircraft-based coordinate system, vary linearly with control stick deflection. These linear and angular velocities are projected onto earth-based reference axes. The earth-reference position, roll, pitch, and heading angles are updated 10 times per second from the corresponding derivatives, using a time interval equal to the update period (.1 second).

Data samples are collected twice per second. Several things are recorded in each sample: values of instantaneous speed, heading, attitude, and position of the aircraft; the values of control stick inputs; status of the autopilot (on or off); elapsed time; and the status of each of the subsystem indicator dials. The pilot's keyboard responses to the subsystem indicator dials are also recorded. These data are stored on disk for later analysis.

The simulation is written in BLISS-11 for a PDP-11/40 minicomputer. The 11/40 is interfaced to a CRT graphics display. The computer also has a disk unit, floating-point arithmetic hardware, and a multiplexed A/D converter which interfaces with the control sticks used by the pilot. The simulation program, along with associated programs such as graphics routines, requires 28K (16 bit) words of core for operation. Floating-point arithmetic is used in simulating the aircraft dynamics; integer arithmetic is used elsewhere for speed.

IV. Experiments

The first experiment will be run using the simulation control and monitoring tasks as described earlier. No computer aid will be present; the pilot will have full responsibility for flying the airplane and monitoring the subsystem indicators. Two variables will be manipulated in these experiments: the complexity of the control (flying) task, and the workload involved with the monitoring task.

Two levels of difficulty of the control task will be used. One level will be provided by normal en-route flight. A more difficult level will be produced by simulated terminal area "landing" maneuvers, where more attention must be given to steering the airplane over a more complex route.

The workload involved with the monitoring task will be manipulated by changing the rate at which subsystem malfunctions occur. The subsystems malfunction in a random-appearing manner. (However, the malfunctions are actually prearranged.) By increasing the rate at which the malfunctions occur, the workload presented to the pilot is increased. Three levels of arrival rate of malfunctions will be used.

From the data collected in these experiments, the response time to subsystem malfunction, the time to service the malfunction, and the control performance of the pilot can be obtained. The probability of a false alarm (pilot mistakes a functioning subsystem for a malfunctioning one), of a missed event (the pilot fails to detect a malfunction when it occurs), and of an incorrect action (the pilot makes a mistake in searching the checklist), can also be estimated. From the results of this experiment, the parameters of unaided pilot decision-making will be determined.

A second experiment, designed on the basis of the experimental results anticipated above, will employ a simple, non-adaptive computer aid. The computer will use a static (i.e., situation independent) policy to determine when it is appropriate to aid the pilot. This experiment will be similar to the first experiment in design. The effectiveness of the non-adaptive computer aid can then be evaluated. Also, within this experiment, we will consider the pilot's ability to detect that the computer aiding system has degraded in performance. In the next section, Chu will discuss approaches to deciding when the computer should aid the pilot.

APPROACHES TO ADAPTIVE COMPUTER AIDING

I. The problem

Considering the allocation of decision making responsibility in a pilot-computer decision making system, our previous simulation analyses (Rouse [75,76]) have shown that overall system performance can be enhanced if the computer's decision making strategy adapts to the current state of the aircraft, the pilot and the computer. In this section of this report, we will discuss results from the literature that are potentially of use in the design of an adaptive system.

The overall pilot-computer system can be characterized as follows:

1) The two decision makers (pilot and computer) are faced with a multi-task event detection and action implementation situation in which a set of monitoring and control tasks randomly demands attention.

2) The decision making situation can be represented as a priority queueing system with two very different and imperfect servers and with imperfect communication between them.

3) To enhance overall system performance, the allocation of decision making responsibility should be dynamic (i.e. situation dependent) while conflict between the two decision makers should be avoided.

Given the decision making situation characterized in the above statements, several questions arise. These are concerned with defining and quantifying the state of the pilot-computer-aircraft system, and then, assuming the computer to be a copilot of sorts, determining how to schedule computer decision making (e.g., when to turn it on or off) as a function of system state.

As noted earlier, we have chosen to view the pilot-computer decision making situation as a queueing problem. Thus, the numerous results which we will review here come mainly from the queueing literature. However, before pursuing this review, we want to mention an alternative view of the decision making situation. Namely, the situation might be described as a stochastic control process with a controller structure adjusted dynamically for on-line estimation and control in an uncertain system environment. Such a representation of our system would require a knowledge of the underlying sequential decision models of each controller (i.e. each decision maker). In this control context, we might start with the control of linear systems subject to random disturbances as well as random system failures and the detection and correction of failures. Thus it is possible to model the pilot-computer decision making situation via the stochastic differential equation approach. However, to avoid going into the microscopic details of system state, we have chosen to view the problem as optimal control of a queueing discipline. The remainder of this section of this report reviews the literature appropriate to this approach.

II. Review of the literature

II-A. Introduction

The early works on the control of queues were essentially descriptive analyses of a set of plausible control policies from which "optimal policies" were selected by mathematical optimization techniques. More recently, researchers have begun to employ a Markovian decision model to solve queueing control problems. The state of the queueing system (e.g. the numbers of customers in the system, the status of service facilities) is observed at certain times (the decision epochs), a control action is taken and costs (e.g. entering prices, waiting costs, service costs, switching costs, servicing rewards) are incurred. The aim is to find a policy, a strategy for choosing successive actions, that will minimize the expected total system cost (average or discounted) over a specified time horizon (finite or infinite). The typical solution approach is first to show the existence of an optimal stationary policy then to design an algorithm employing techniques such as dynamic programming, linear programming, policy improvement or value iteration for computing an optimal policy. While it would be interesting to pursue discussion of these techniques with full analytical detail, instead we shall concentrate on summarizing the models that have been used and the results that have been obtained.

It is usually convenient to categorize the system of interest by 1) the system structure (e.g. M/M/s, M/G/1, GI/M/1), 2) the decision variables (e.g. service parameters (service rates, number of servers, etc.), arrival parameters, queueing disciplines, operating time parameters). In the following section static models are discussed before dynamic ones since many dynamic control models cannot be solved separately from the static design models. Also, since quite a few of the server control models have served as bases for the development of corresponding priority control models, we will start section II-B with discussions of server control (mainly with server on-off) followed by discussion of service rate control.

In section II-C we will discuss many interesting models for the control of arrivals. Section II-D considers various priority control schemes which include a whole class of service control-based priority queueing systems and a couple of examples on dynamic priority assignment policies. Section II-E discusses interactions in decision and control environments and those in sequential decision and stochastic control contexts.

II-B-1. Control of server

Here the control action is to turn the server on or off at the service completions or at the customer arrivals when system is empty. Costs are charged for 1) switching (on-off), 2) servicing and 3) holding the customers.

Blackburn [72] considers the problem of controlling a M/G/1 queue by turning the server on and off. He shows that the optimal stationary policy which maximizes expected discounted reward over an infinite horizon has a simple critical number characterization : (M,m). This (M,m) policy is to provide no service if the system size is m or less, and to turn the server on when the size is greater than M. This result is quite similar to those obtained from inventory theory. Blackburn analyzes the problem as a Markov renewal decision process with constant switching costs and linear holding costs and shows that the m will be either 0 (that is to turn service off when the system is empty) or -1 (the server is never off). His model allows for customer balking and reneging.

For the M/G/1 queue with server on-off cost, server running cost and customer waiting cost, results based on the analysis of continuous time and discrete time by Heyman [68] and Magazine [71] also yield (M,m) as an optimal solution. Deb [76] extends the results to a situation of a batch service of size up to Q, the control limits (M,m) are computed for M/M/1 case with linear service time costs.

Considering other than the Markovian decision point of view, Man [73] and Shaw and Hsuan [74] have separately considered the stochastic optimal control strategies for arrival rate regulated and service rate controlled systems with time varying input traffic demand. A set of continuous time state-space differential equations is derived, the maximal principle is applied and a two point boundary value problem is obtained. However the complexity in computation and implementation makes the approach formidable. Based on the arguments of the Bang-Bang control solution, the authors suggest that the "fixed threshold" policy would be an acceptable suboptimal solution.

Shaw [72] in his work on optimal ramp control found an optimal on-off type two-level threshold controller with respect to minimizing the sum of queue length and customer rejection. In his later work [76], he further shows that the optimal customer rejection (diversion) to minimize total delay of the accepted customers and the rejected customers is of the control-limit form for the M/G/1 case. An explicit expression for the optimal threshold for the M/M/1 case is given as either 0 or infinity depending on the traffic intensity and the relative importance of waiting loss and downtime loss in the maintenance-repairing context.

Balachandran [73] has considered the same on-off policy with control measures determined by the unfinished work in the system (hence the approach requires that service time be known for customers in the system). This so-called "D-policy" is later proven by Balachandran [75] to be superior to the usual N-policies (which turns the sever on when the queue size reacher the value N) for exponential service distribution with decreasing failure rates.

II-B-2. Controls on service rates

In this case a service rate can be chosen from a set of the allowable service rates at customer arrivals or at service completions.

Crabill [72] has employed a continuous time Markovian decision model to investigate the queueing situation with a finite number of exponential service rates μ_i each with associated cost rate r_i. An optimal long run average cost rate (including holding and service costs) stationary policy is such that $\mu_i > \mu_j$ for $i > j$, where μ_i is the optimal service rate when i customers are in the queue. The scheme is primarily a set of sufficient conditions for eliminating from consideration certain service rates and conditions for the existence of an optimal stationary policy. This agrees with the intuition that under certain cost conditions it is optimal to always use the fastest service rate. Crabill [74] extends the previous model to a maintenance system considering costs depending on service rates and costs due to lost production.

Lippman [75] generalizes previous results by implementing a cost structure which consists of a customer holding cost L, a reward for service completion and a service cost rate which increases with the service rate employed. He shows that the optimal service rate (when the discount factor is Z, when there are I customers in the system and N transitions remaining) is an increasing function of I, N, and L/Z. To conclude this section, Crabill [72,74] shows that there exists a monotone optimal average cost policy and extends the result to include switching costs and production losses. Lippman [75] establishes the existence of monotone optimal discounted and average cost policies.

II-C. Control of arrivals

There are many contexts, such as the admission scheduling and routing in a hospital, in which the service rate is not the principal operational problem. Quite recently the optimal arrival control models have been developed in a rather ad-hoc way. In this section we will discuss four scattered categories on control of arrials with the first two categories closely related to the priority assignment models discussed in the next section.

The first category is an extreme control, accepting or rejecting each customer. This amounts to changing the arrival rate from its normal level (accept) to level zero (reject). The second category is based on varying the arrival rates in an intermediate manner. The third category is based on optimal queue formation. The forth category is that in which the customers, rather than the queue, make the decision of accepting or refusing entry to the queue.

Not allowing waiting lines or queues (in the first category - the extreme control), Keilson [70] has considered an M/G/1 queue with arrivals accepted or rejected. Each type j customer entering the system offers a reward R_j for being served and a service time distribution with mean rate μ_j. Observing these along with the number N of customers in the system, the arrival can either be accepted or rejected. Cost are charged for holding the customers in the system. The optimal average (and discount) cost policy over infinite horizon is found to be: Accept a type j customer if and only if $\mu_j R_j > Q$, where $Q > 0$ is easy to find. Various versions of this problem have been analyzed by others with the result that certain μ-R relations lead to similar optimal monotone policies. In the next section, we will also discuss stable customer payment policies for purchasing a priority assignment in a queueing system (Balachandran [72]).

In the second category (the intermediate arrival rate control), for the problem of an M/M/s queue with variable arrival rates and finite waiting places, Low [74] shows the existence of an optimal monotone average cost policy, i.e. an optimal arrival rate is a nonincreasing function of the number in the system. He also gives an algorithm for computing this policy. Kakalaic [69] has considered an M/G/1 queue of J streams of customers with arrival rate chosen from the interval [0,a]. For each class there is a rejection cost, an acceptance reward and a holding cost. The author finds that the optimal policy for optimal average cost is of the bang-bang type; that is to accept all customers of class i (rate a) or accept none (rate 0). This indicates that the optimal policies of the first category satisfy a more robust set of conditions than originally specified.

In the third category (the optimal queue formation), Nazarov [76] considered the G/M/2 queue with limited waiting places. He proposed an optimal dynamic (state-dependent) strategy of assigning arrivals to a service channel for a given functional cost structure which includes cost of holding customers and cost for loss of customers. The decision is determined by the function d(i,j), the conditional probability of assigning the arrival to the first facility given that the system has i customers in channel 1 and j customers in channel 2. Modeled as a linear Markov process and approached by the dynamic programming and finite difference equation technique, the final results are obtained through a successive approximation to minimize the expected cost function (in a limiting average form). The implementation on a digital computer shows that dynamic formation of the queue, as compared to the conventional static approaches, results in a cost saving of up to 30%. The author suggests the use of the dynamic strategy for loading factor (utilization) near unity. For a total loading factor considerably different from unity the gain is insignificant.

Considering the question of customer control over entering a queue with holding cost c or refusing to enter the queue with lost-service cost r (in the fourth category - the arrival joining/rejection), Noar [69] gives optimal joining rules for a M/M/1 queue: join the queue if and only if the queue size N is less than or equal to M for self optimization (customer himself) and m for social optimization (customer + system) where in general $M > m > 0$. Thus self optimization leads to a more congested system than that of social optimization, as we might have expected.

II-D. Control of queue discipline - Priority models

The separation of customers at a service facility into distinct classes, and then servicing the classes according to the measure of importance gives rise to what is known as a priority queue. Jaiswal [68] gives an excellent discussion of the basic structure of priority queues. As Bronshteyn and Rykov [65] have pointed out, systems with optimal series of priorities are better than those without priority assignments, although the latter is better than the system with priorities inappropriately established. To optimize a priority queue, we might want to determine the assignments of customers to priority classes, or given a natural set of classes to determine the service discipline so as to optimize some measure of system effectiveness.

Most of the priority control papers have evolved as an outgrowth of classical descriptive models which are static (exogenous or state-independent) and deterministic (in the decision criterion). Several of the arrival control approaches to priority assignment given in the previous section fall into this category and they will be discussed in the last part of this section. First we will examine the classical results of Cox and Smith [61], developing the models in parallel with the generalizations on dynamic scheduling disciplines. Secondly, we will discuss several dynamic priority models in the context of time-dependent or state-dependent categories, and thirdly, the control of server disciplines in priority queues.

In an M/G/1 nonpreemptive priority queue of finite J flows of customers with waiting cost c_j and service rate μ_j, Cox and Smith [61] have obtained an optimal priority assignment policy (called the μ-c solution) which states: of all the possible nonpreemptive work-conserving stationary queueing policies, the head-of-the-line discipline with the highest priority assigned to the class j customers with higher $\mu_j c_j$ products is that which minimizes the average waiting cost. This result has been utilized in many simple priority models. Among other extensions of it is the Bronshteyn's [65] generalization of μ-c solution to the preemptive discipline for an M/M/1 queue. Practically speaking the μ-c solution can be employed as an operating rule for priority assignment by implementing a partitioning scheme which includes a set of J+1 critical values. Considering the same M/G/1 head-of-the-line discipline with a finite number of priority classes, Beja and Sid [75] present two properties of an optimal priority partition:

1) The μ-c products are monotonic on a set of optimally partitioned intervals. 2) It is optimal to assign strictly as many different priorities as are allowed. The authors further discuss the situations for different information sets on μ and c. If μ or c is unobservable then, instead of μ-c products, the estimates $c/E[t_s \mid C=c]$ or $\mu\ E[c \mid U=\mu]$ is used for priority assignment.

A summary of results for application of the extended μ-c solution to dynamic scheduling of a two-class M/G/1 queue is given by Harrison [76]. He shows that there is a modified static policy that is optimal in maximizing the expected service reward r_j minus holding cost c_j incurred over an infinite time horizon. The modified optimal policies are: 1) If $c_j=0$ then, instead of μ-c, we have μ-r policy., 2) If $c_1 > 0$, $c_2 > 0$, λ_2/μ_2 (the traffic intensity of flow-2) > 1 and $\mu_2 c_2 > \mu_1 c_1$ then serve class 2 customers or idle. 3) If $c_2 > c_1=0$ and $c_2 > (r_1(1-\lambda_2/\mu_2))/(2E(t_{s_2})\lambda_2)$ then serve class 2 customers or idle. 4) If $\lambda_2/\mu_2=1$ or $\mu_1 c_1=\mu_2 c_2$ then the μ-c solution does not apply to these degenerate cases; criteria based on the higher moments of service time distributions are required.

There are quite a few papers that include discussions of dynamic priorities. Jackson [60] is one of the first to introduce the concept of the time-dependent priorities. Assuming that each class i customer is assigned on arrival a random urgency number a_i, Jackson's [65] dynamic priority for this customer is given by a random variable $p_i(t) = a_i + b(t - t_i)$, where b is a constant common to all customers (namely 1) and t_i is the time of arrival. So a newcomer takes precedence over one in the queue if and only if the difference between the former's urgency number and that of the latter is larger than the time the latter has spent waiting. Actually we have a fixed priority type of queue where a customer's priority does not change relative to the other customers in the queue when he arrives. Under certain hypotheses, the author gives bounds for expected waiting times. Kleinrock's [67] time-dependent priority model assumes a proportionality between priority and waiting time for each customer: $p_i(t) = b_i(t-t_i)$. Thus, Kleinrock's model allows changes in relative priority.

Another category of dynamic priority concepts was introduced by Vekierov [67] with the implication of state-dependent priorities. Rykov and Lembert [67] show for M/G/1 queues that the optimal dynamic priorities are simply the ordinary or static priorities. In another words, the μ-c solution is optimal in an essentially broader class of dynamic service policies. However this nice property will not hold for a system with a finite queue length, as we will see next.

Mova and Ponomarenko [74] considered an M/M/c queue with finite r_j waiting places. There are n different flows of customers with Poisson arrival rates λ_i and c servers with same service rate μ_i. Denoting the system state for which there are i_j class j customers waiting for service by $(i_1,i_2,\ldots,i_n)$, the service policy $d_j(i_1,i_2,\ldots,i_n)$ is to choose a class j customer with probability $d_j(i_1,i_2,\ldots,i_n)$. The authors set up the state probability equations (a Markov process, stationary regime), apply the linear programming and necessary condition for the principle of optimality to obtain an optimal ordering for servicing. The optimal d for a cost structure that includes waiting and lost customer costs in the state penalty form are found to be either 0 or 1, or more precisely,

$$d_j(i_1,i_2,\ldots,i_n) = \begin{matrix} 0 & j \neq s \\ 1 & j = s \end{matrix} \quad ,$$

where $s = s(i_1, i_2, \ldots, i_n)$ is the state index indicating the system's current configuration. Determination of the s's involves solving a set of quasilinear equations. An example for an M/M/1/2 case is given. Under the assumption that $r_1 = r_2 = 1$; $\lambda_1 = 0.3$, $\lambda_2 = 0.5$, $\mu=1$ the minimum cost service discipline obtained is to serve class 1 flow when $(i_1,i_2) = (1,0), (2,0), (2,1)$ and to serve class 2 flow when $(i_1,i_2) = (0,1), (0,2), (1,1), (1,2), (2,2)$. The results yield a benefit of a 4.25% decrease in cost compared to fixed priority assignment to class 2 flow. Thus, in this case the classical rule of assigning relative priorities (such as the μ-c solution) is not optimal.

It seems natural to expect that the success realized for control of the server (section II-B-1) might be extended to a priority queueing model. For the 2 priority classes M/G/1 queue with server on-off at the arrival-departure epochs, Bell [71,73] has proven the existence of an optimal average cost policy of the (a,b,c) type. This optimal strategy is to turn the server off only when the system is empty and to turn the server on the first time that $a \cdot n_1 + b \cdot n_2 > c$ where a,b,c are nonnegative constants with $a+b > 0$ while n_1 and n_2 are the numbers of class 1 customers and class 2 customers in the system. For the general m priority classes, the optimal control actions are simply characterized by the m-dimensional hyperplane of the form: $a_1 \cdot n_1 + a_2 \cdot n_2 + \ldots + a_m \cdot n_m = c$. The results hold for a general cost structure which includes holding, running and switching costs.

For the two priority classes situation mentioned above, Tijms [74] derives an expression for the average number of the class i customers and determines the best (1,1,c)-, (1,0,c)- and (0,1,c)- policies with respect to an average linear cost criteria. These efforts are mainly an attempt to fit the control of server model into the priority class situations and the (a,b,c) policies can be viewed as the counterparts of the (M,m) policies of Heyman [68].

Not satisfied with the "static" design of a priority rule, Bell [73] presents the optimal priority policies that allow for changing priority rules as a function of the system's current state. He considers an M/G/1 $n=2$ case with holding costs $h_1 > h_2$ and the service option of FCFS (first come first serve) or promotion of class 1 customers further back in the queue with penalty R. Bell shows the existence of an ϵ-optimal minimum average cost stationary policy. The optimal policy is: Serve any class 1 customers at the head of the line or to promote the last class 1 customer and to serve him next if a class 2 customer is at the head of the line and there is a class 1 customer in the queue with at least K class 2 customers in front of him. The K is determined by the relation $K > (R\mu)/(h_1-h_2)$. The results can be extended to the case of different service time distributions between the classes. However if the penalty R is state dependent, which unfortunately is usually the case, the optimal decision in any state will typically depend on the precise waiting list of class 1 and class 2 customers in the state description.

Balachandran [72] and Kleinrock [67] have considered models in which a customer's priority is assigned according to the amount he is willing to pay, higher payment receiving higher priority (Balachandran calls it "purchasing priority", Kleinrock [76] calls it "bribery priority"). Balachandran develops methods to determine the conditions for stable payment policies in situations where optimal policies are not stable. He defines the stable policy as a policy under which no individual customer can improve his situation by deviating from the stable policy, given that all other customers use the stable policy. He illustrates that an optimal average cost policy need not be stable. For M/M/1 queues he exhibits two rules which are stable under certain conditions. He also considers the case for which waiting time costs increase with time.

Nair and Neuts [71] adopt a probabilistic approach to compare waiting times under three priority rules: FCFS, SPT (shortest processing time first) and LPT (longest processing time first). The analysis relies on the theory of semi-Markov and renewal processes. The comparison is achieved essentially by the derivation of the Laplace-Stieltjes transform of the trivariate distribution of transient and steady state waiting times under each of the three rules. This approach makes it possible to answer the questions like: What is the probability that the waiting time (as a random variable here) at t of a customer under SPT rule whose service time is x is less than that of a customer under FCFS. However the complicated results shown in this paper could be a major drawback as far as the implementation on an adaptive system is concerned.

II-E. Miscellaneous Issues - On design in decision and control environments

As has been shown in this survey of the optimal design of queueing systems, the queueing approach turns out to be a very handy and powerful tool for modeling a decision and control system. There are several important issues concerning the interaction and interfacing between the queueing model and design environment that have to be coordinated or resolved before we can talk about the application and implementation of this powerful methodology.

In the preemptive situation the SRPT (shortest remaining processing time first) disciplines are employed in substitution for the SPT for the system to be optimal.

Thus far we have been mostly concerned with optimal queueing systems operating over an infinite horizon with a homogeneous evolution of time. Another class that arises in a non-stationary environment is concerned with how to best operate the system over time with a given cost structure. Prabhu [74] has considered an M/G/1/FCFS system, employed the optimal stopping time concept and derived the optimal stochastic control theory. The optimal control is achieved by choosing the stopping time T that minimizes the overall cost over a given class of stopping times. Hsuan and Shaw [76] discuss the optimal dynamic control and repairman assignment policies for a linear stochastic system with Markov jump parameters. They formulate the problem as a Markov decision process and present an elimination of equivalent policy technique which, incorporated with the control limit techniques developed earlier, makes their approach a little more promising. Man [73] derives an optimal control strategy for a M/M/s/N system with arrival rates controlled. The control system is dynamic and is capable of handling the time-varying input traffic demand in a way to employ the on-line traffic load measurement for updating the new optimal control policy. The system is therefore claimed to be less vulnerable to the uncertainties present in the system parameters.

III. References

[1] Bagchi, T. P. and Cunningham, A. M. (1972) "Bayesian approach to the design of queueing systems", INFOR 10-1, pp. 36-46.

[2] Balachandran, K. R. (1970) "Parametric priority rules: an approach to priority queues", Oper. Res. 18, pp. 426-40.

[3] Balachandran, K. B. (1972) "Purchasing priorities in queues", Management Science 18-5, pp. 319-26.

[4] Balachandran, K. B. (1973) "Control policies for a single service system", Management Science 19-9, pp. 1013-18.

[5] Balachandran, K. B. (1975) "On the D-policy for the M/G/1 queue", Management Science 21-9, pp. 1073-76.

[6] Bell, C. E. (1971) "Characterization and computation of optimal policies for operating an M/G/1 queueing system with removable server" Oper. Res. 19-1, pp. 208-18.

[7] Bell, C. E. (1973) "Efficient operation of optional-priority queueing systems", Oper. Res. 21-3, pp. 777-86.

[8] Bell, C. E. (1973) "Optimal operation of an M/G/1 priority queue with removable server", Oper. Res. 21-6, pp. 1281-90.

[9] Beja, A. and Sid, E. (1975) "Optimal priority assignment with homogeneous waiting costs", Oper. Res. 23-1, pp. 107-17.

[10] Bertsekas, D. P. (1976) "Dynamic programming and stochastic control", New York: Academic Press.

[11] Blackburn, J. D. (1972) "Optimal control of a single server queue with balking and reneging", Management Science 19-3, pp. 279-313.

[12] Bronshteyn, O. I. and Rykov, V. V. (1965) "Optimal priorities in queueing systems", Eng. Cybernetic 6, pp. 26-35.

[13] Carbonell, J. R. (1966) "A queueing model of many instrument visual sampling", IEEE Tran. HFE 7-4, pp. 157-64.

[14] Chan, W. and Chung, W. K. (1971) "Computer controlled queueing systems with feedback", Proc. IEE 118-10, pp. 1373-77.

[15] Chung, W. K. (1974) "Computer-controlled queueing system with feedback and swap times", Proc. IEE 121-10, pp. 1080-85.

[16] Cox, D. R. and Smith, W. L. (1961) "Queues", London: Methuen.

[17] Crabill, T. B. (1972) "Optimal control of a service facility with variable exponential service times and constant arrival rates", Management Science 18-9, pp. 560-66.

[18] Crabill, T. B., Gross D. and Magazine, M. J. (1973) "A survey of research on optimal design and control of queues", Report, The George Washington University.

[19] Crabill, T. B. (1974) "Optimal control of a maintenance system with variable service rates", Oper. Res. 22, pp. 736-45.

[20] Deb, R. K. (1976) "Optimal control of batch service queues with switching costs", Adv. Appl. Prob., March 1976.

[21] Harrison, J. M. (1975) "Dynamic scheduling of a multiclass queue: discount optimality", Oper. Res. 23-3, pp. 270-82.

[22] Harrison, J. M. (1975) "A priority queue with discounted linear costs" Oper. Res. 23-2, pp. 260-79.

[23] Harrison, J. M. (1976) "Dynamic scheduling of a two-class queue: small interest rates", SIAM App. Math. 31-1, pp. 51-61.

[24] Heyman, D. P. (1968) "Optimal operating policy for M/G/1 queueing systems", Oper. Res. 16, pp. 362-82.

[25] Hsuan, J. and Shaw, L. (1976) "Optimal control and synamic repairman assignment for a linear stochastic system", ORSA/TIMS March 1976, Philadelphia.

[26] Jackson, J. R. (1960) "Some problems in queueing dynamic priorities", Naval Res. Log. Quart. 7, pp. 255-69.

[27] Jackson, J. R. (1965) "Queue with dynamic priority discipline", Management Science 8, pp. 18-34.

[28] Jaiswal, (1968) "Priority queues", New York: Academic Press.

[29] Kakalic, J. (1969) "Optimal dynamic operating policies for a service facility", Report, OR Center MIT.

[30] Keilson, J. (1970) "A simple algorithm for contract acceptance", Opsearch 7, pp. 157-66.

[31] Kleinrock, L. (1967) "Optimal bribing for queueing positions", Oper. Res. 15, pp. 304-18.

[32] Kleinrock, L. (1976) "Queueing systems, Vol. 2", New York: Academic Press.

[33] Lippman, S. A. (1975) "Optimal control of an M/M/1 queue with variable service rates", Bull. Oper. Res. Soc. Amer. 23-b419, Fall 1975.

[34] Low, D. (1974) "Optimal dynamic pricing policies for an M/M/s queue", Oper. Res. 22, pp. 545-61.

[35] Magazine, M. J. (1971) "Optimal control of multichannel service systems", Naval Res. Logist. Quart. 18, pp. 177-83.

[36] Man, F. T. (1973) "Optimal control of time varying queue", Management Science 19-11, pp. 1249-56.

[37] Mova, V. V. and Ponomarenko, L. A. (1974) "On the optimal assignment of priorities depending on the state of a servicing system with a finite number of waiting places", Tran. Eng. Cybernetics 12-5, pp. 66-72.

[38] Nair, S. S. and Neuts, M. F. (1971) "An exact comparison of the waiting times under three priority rules", Oper. Res. 19-2, pp. 414-23.

[39] Naor, P. (1969) "The regulation of queue size by levying tolls", Econometrics 27, pp. 15-23.

[40] Nazarov, A. A. (1976) "Optimal formation of queues in a multichannel queueing system", Auto. and Remote Control, pp. 1242-44.

[41] Prabhu, N. U. (1974) "Stochastic control of queueing systems", Naval Res. Logistic Quart., Sep. 1974.

[42] Reed, F. C. (1975) "A difference equation approach to the optimal control of a multiclass queue with discounted cost", TR NWC-TP-5596 Naval Weapons Center 1975.

[43] Rouse, W. B. (1975) "Human interaction with an intelligent computer in multi-task situations", Proc. of the 11th Annual Conf. on Manual Control, NASA - Ames Research Center, pp. 130-43. May 1975.

[44] Rouse, W. B. (1976) "Adaptive allocation of decision making responsibility between supervisor and computer", Proc. of the Internat. Symposium on Monitoring Behavior and Supervisory Control, Berchtesgaden, Germany, March 1976.

[45] Rykov, V. V. and Lemberg, E. Y. (1967) "Optimal dynamic priorities in single line queueing systems", Tran. Eng. Cybernetics 5-1, pp. 21-30.

[46] Schrage, L. (1975) "Random results in scheduling: implications of queueing theory for scheduling", Working Paper, University of Chicago.

[47] Shaw, L. (1972) "On optimal ramp control of traffic jam queues", IEEE Tran. on Automatic Control 17-5, pp. 630-37.

[48] Shaw, L. and Hsuan, J. (1974) "Optimal control and dynamic repairman assignment for stochastic linear systems", Report, Polytechnic Inst. of New York.

[49] Shaw, L. (1976) "Busy period control of queues based on waiting time at arrivals", Oper. Res. 24-3, pp. 543-63.

[50] Stidham, S. and Prabhu, N. U. (1974) "Optimal control of queueing systems", Math. Method in queueing systems, Springer Verlag 1974.

[51] Tijms, H. (1974) "A control policy for a priority queue with removable server", Oper. Res. 22, pp. 833-37.

[52] Vekierov, Y. B. (1967) "Optimal absolute dynamic priorities in queueing systems", Tran. Eng. Cybernetics 5-2, pp. 87-91.

HUMAN DECISION MAKING IN MULTI-TASK SITUATIONS*

I. The problem

There are many situations in which the human decision maker is responsible for several tasks and must decide which of these tasks is to be performed. Industrial monitoring, air traffic control, aircraft piloting, and investment decision making are examples of such multi-task situations. As the complexity and performance demands of such situations increase, the human is faced with more tasks of greater variety. He may be unable to maintain acceptable performance with this increased workload.

Attempts to reduce the human's workload and make more efficient use of his cognitive abilities are resulting in changes in the human's role in the operation of complex systems. He is no longer an integral part of many of the control loops within the system and instead devotes much of his capacity to monitoring machine controlled processes for events which require diversion of his attention to a particular process. In an air traffic control situation, for example, he may monitor the positions of several planes and divert his attention to a particular plane when it strays from its assigned path or begins its landing approach.

A problem faced by the human in this situation is that the diversion of his attention to a particular task may cause him to miss the occurrence of events that relate to his other tasks. This problem might be alleviated by the addition of a computer aid with some event detection capability.

---

* This section is based on the Ph.D. thesis proposal of J. S. Greenstein entitled "Human Decision Making in Multi-Task Situations: Event Detection, Attention Allocation and Implications for Computer Aiding." This research is partially supported by the Joint Services Electronics Program.

If an interesting event were detected while the human's attention was diverted to another task, the aiding system could notify the human, or, if the aid had sufficient ability, it might make decisions without notifying the human.

The use of a computer aid with responsibilities overlapping those of the human lessens the risk involved in the human's diversion of attention to one task. But such a responsibility overlap introduces the question of how decision making conflicts are to be avoided. A set of simulation studies by Rouse [75,76c] considered a situation in which human and computer monitored and acted with respect to a set of processes. It was found that the level of interaction in terms of what the human and computer knew about each other's actions had a significant effect on the number of conflicts and on overall performance. As feedback between the human and computer about each other's actions and planned actions was increased, conflict was reduced. An improvement in overall system performance was the result.

A design question that arises, then, is that of devising some method of letting the human and computer know what each other is doing. Computer-to-human communication does not appear to be a serious problem. A simple status indicator could be used to inform the human of the tasks to which the computer is attending.

Human-to-computer communication presents greater difficulties. If the human must consciously keep the computer apprised of his present and planned actions he may well find himself spending more time communicating his state than carrying out his tasks. The computer aid would not be doing much to reduce the human's workload in such a case. Instead, the computer must have some way of inferring the human's state. That is, the computer needs some model of how the human detects events and allocates his attention among tasks. With this model the computer can determine its actions so as to minimize conflicts with the human.

There is a need, then, for models of the human as an event detector and attention allocator in multi-task situations. These models would be of use both in the design of computer aiding systems and in their implementation. The insights of the human decision making process the models provide would also be of use in the analysis and design of other systems in which the human has a decision making role. Some preliminary approaches to modeling the human in event detection and attention allocation tasks have been proposed by Rouse and Greenstein [76b] and will be discussed later in this paper. This research proposes to extend and validate these basic models. Specifically, it proposes to conduct an experimental investigation of human performance in event detection and attention allocation in a multi-process monitoring situation. The results of these experiments will then be used to validate and extend the preliminary models of the human in such tasks. Finally, the implications of these models to the design and implementation of a computer aiding system will be considered.

II. Related research

Miller and Elkind [67] and Gai and Curry [75,76] have modeled the human in event detection tasks. Miller and Elkind consider a situation in which the human is a part of the control loop in a single task situation. They employ the idea that the human has an internal model of the controlled process which he uses to predict the effect of his control movements.

The detection of changes in process dynamics is based on the difference between the actual effect of his control movements and the effect his internal model predicts. Gai and Curry consider a situation in which the human is outside the control loop and monitors a process for failures. Their model of the human employs estimation theory and decision theory and assumes that the monitored process can be represented as a Gaussian process. Failures are represented by a change in the mean of the monitored process.

Several researchers have modeled the manner in which the human allocates his attention in multi-process monitoring tasks. Senders [64] employs an information theory approach while Carbonell [Carbonell, 66; Carbonell, Ward, and Senders, 68] and Senders and Posner [76] suggest queueing theory approaches. Decision theory [Kvalseth, 75] and optimal control theory [Curry, Kleinman, and Hoffman, 75; Kleinman and Curry, 76] have also been used to model the human in monitoring tasks. Smallwood [67] uses the internal model concept and proposes that the human relies upon internal models of the processes he is monitoring in making his attention allocation decisions.

All of these attention allocation models emphasize task monitoring. Their primary intent is to provide predictions of performance and workload for specific display designs. The attention allocation problem presented in this paper is different in that it is less concerned with the allocation of attention among monitoring tasks than with the allocation of attention between task monitoring and action with respect to events observed during task monitoring. Action with respect to an event is assumed to require the human to divert his attention from task monitoring for some period of time. It is also assumed that all of the tasks can be monitored simultaneously in the sense that the human updates his knowledge of all tasks whenever he chooses to monitor.

This proposal suggests that models of human decision making might be used in computer aiding systems to reduce the human's workload in multi-task situations. Freedy and his colleagues [Freedy, et. al., 71; Freedy, Weltman, and Lyman, 72; Freedy, Weisbrod, and Weltman, 73; Weisbrod, Davis, and Freedy, 75; Steeb and Freedy, 76] have developed adaptive computer aiding systems for single task situations in which the computer aid learns to perform a task by observing the human's performance in the task. When sufficiently confident of its ability, the computer assumes control of the task from the human. Their research is directed toward developing human factors criteria for such shared decision making. Also concerned with the development of human factors criteria for the design of computer aiding systems are the experimental investigations of Poulton and his colleagues [74]. They study the effects of various methods of computer assistance on human performance in sonar signal detection and classification tasks.

Rouse [75,76b,76c] has investigated the issues arising when human and computer interact in multi-task situations. He considers situations in which the human and computer have overlapping abilities and responsibilities and shows, among other things, that the amount that the human and computer know about each other's actions has a significant effect on overall system performance. It is the intent of the research proposed in this paper to facilitate human-to-computer communication in such multi-task situations through the development of models of the human's performance in such tasks.

III. Proposed approach

There are several possible approaches to the modeling of the human in event detection tasks. Signal detection theory might be employed [Sheridan and Ferrell, 74] but by itself it is of little use to a computer aid in explaining how the human actually detects events. If the task state can be modeled as a stochastic dynamic system an estimation theory approach [Gai and Curry, 76; Rouse, 76a; Willsky, 75] would be appropriate, but many realistic multi-task situations would not yield easily to such analysis. Discriminant analysis [Tatsuoka, 71; Afifi and Azen, 72] appears to offer advantages as an approach to modeling event detection. It offers the possibility of allowing the computer aid to adaptively define what the human decision maker considers to be an event. It might also allow the costs of errors to be inferred from the human's actions. Thus, a computer aid using a discriminant analysis technique might analyze the human's decisions over a period of time and assist with the event detection tasks when sufficiently confident of its own ability.

The problem of attention allocation in multi-task situations might be modeled as a system in which various tasks randomly demand the attention of the decision maker with various frequencies and for various durations. Additionally, some tasks would be expected to be of greater importance than others. If minimization of the decision maker's delay in acting upon events, particularly those relevant to important tasks, is a reasonable measure of performance, then a queueing theory approach may be useful in modeling the human as an attention allocator.

Some preliminary models of human performance in event detection and attention allocation tasks have been developed using the approaches mentioned above. They are discussed in some detail in an earlier paper [Rouse and Greenstein, 76b]. These models indicate that in event detection tasks the a priori probabilities of event occurrences, the values of correctly responding to events and not responding to non-events, and the costs of making false alarms and missing events are important parameters. In multi-task situations in which the decision to react to an event requires the human to divert his attention to a specific task for some period of time, the decision making process is considerably more complicated. Important parameters in this situation include the a priori probabilities of event occurrences in the different tasks, the amount of time attention must be diverted when responding to events in the different tasks, and the costs of delays in responding to events in the different tasks. Another important parameter in this situation appears to be the human's planning horizon, the number of actions or units of time that the human plans ahead.

A set of two experiments has been planned to validate the proposed models of event detection and attention allocation in a specific situation. The first experiment, recently completed, investigates the human's event detection behavior, while the second experiment focuses on his attention allocation behavior. The situation employed involves the simultaneous monitoring of several dynamic processes for the occurrence of abnormal events, a situation representative, for example, of monitoring tasks in complex industrial plants, advanced aircraft, or air traffic control. The design of the experimental situation is presented in an earlier paper [Rouse and Greenstein, 76a].

Figure 1 illustrates the display observed by subjects in the first experiment. The display depicts the measured values of the outputs of nine processes over 100 sampling intervals. The processes had identical second order system characteristics with zero-mean Gaussian white noise inputs of identical variance. The displayed measurements of the process outputs were corrupted by additive zero-mean Gaussian white noise sequences which normally had identical variance. An abnormal event in a process was defined by an increase in the measurement noise variance such that the signal-to-noise ratio for each measurement following an event occurrence was decreased to 95% of the signal-to-noise ratio of the previous measurement. Thus abnormal events became more and more pronounced with each successive measurement.

After scanning the nine process histories, subjects were given an opportunity to key in the numbers of processes in which they had decided an abnormal event had occurred. They also positioned a cursor to enter their estimates of the times at which the events occurred. Upon completing their responses they were given feedback regarding the actual states of the processes they had keyed in. The display was then erased, current scores were given, any abnormal processes detected were returned to the normal state, and a new display depicting the process histories advanced 10 units in time was generated as illustrated in Figure 2. Dashed vertical lines were used to indicate to subjects when they last responded to each process.

Subjects were allowed to respond to as many events as they thought had occurred. They were awarded points for their hits, receiving

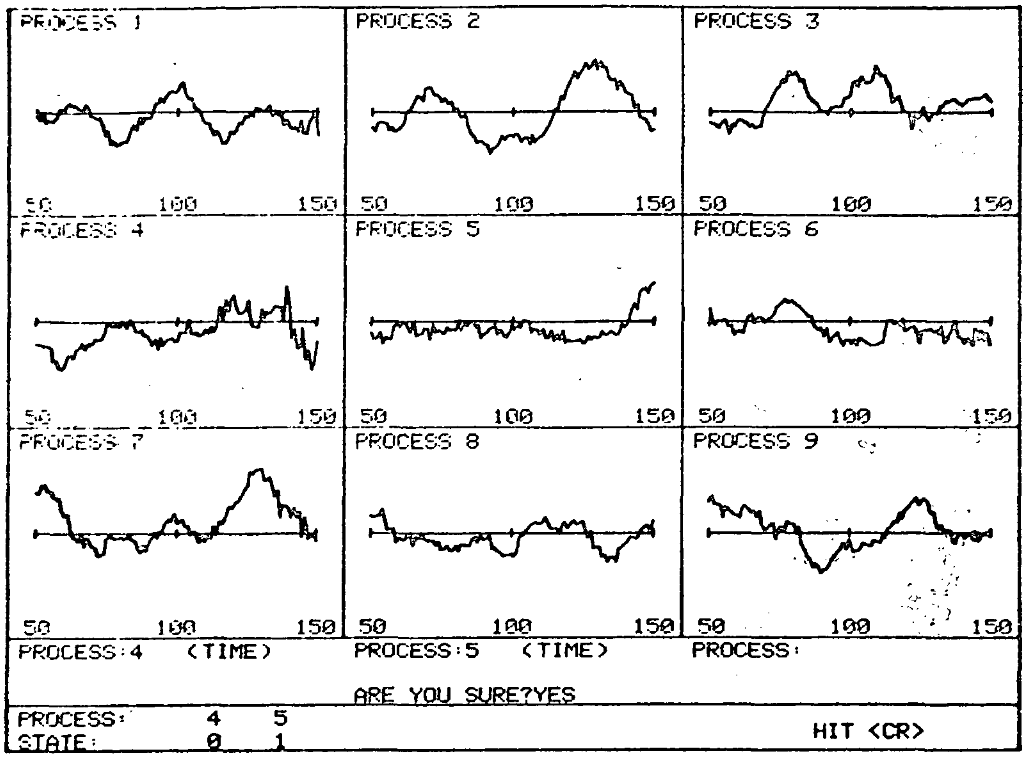

Figure 1: The Multi-Task Situation

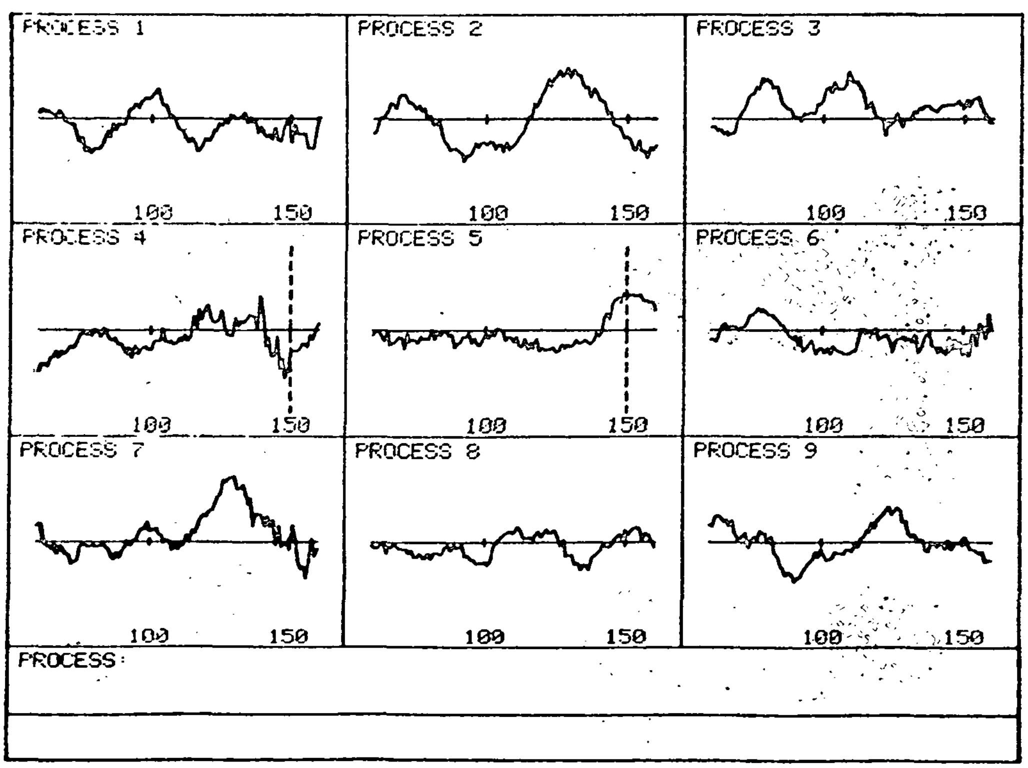

Figure 2: An Updated Display

high scores for responding to events soon after their occurrence and lower scores for tardier responses. A fixed number of points was deducted for each false alarm. Subjects were allowed to study the displays as long as they wished, but any time taken beyond the first minute on each iteration reduced the number of points awarded for hits made on that iteration.

Eight subjects were each given three trials spaced over several days. Each trial was 20 iterations long. The first and third trials given half the subjects were identical and had one event scheduled per iteration. The second trial scheduled the same events as the first and third trials, but also scheduled an additional event each iteration to permit study of the effect of a priori event probabilities on performance. The rest of the subjects were given these runs in different order so that the first and third runs had two events scheduled per iteration while the second run had one per iteration.

The results of this experiment will be used to test the adequacy of a discriminant analysis approach to modeling event detection behavior. The time histories of the subjects' decisions will be analyzed using discriminant analysis techniques to determine model parameters and the resulting model predictions will be compared with the subjects' actual performance. The ability of the model to adapt to individual differences and time varying behavior will also be considered.

A second experiment is planned to investigate the human's attention allocation behavior in a situation similar to the one employed in the first experiment. But whereas in the first experiment subjects were allowed to respond to as many events each iteration as they thought had occurred, in this experiment they can choose one process in which they believe an event has occurred or they can choose none. The amount of time the process histories are advanced on the succeeding iteration is dependent upon the decision they make. The decision not to enter a process number represents a decision to continue monitoring all processes for events. The process histories will then be advanced only 10 units in time on the next iteration. The decision to respond to an event they believe has occurred in one of the processes represents a diversion of attention to that process. The process histories will then be advanced more than 10 units in time on the next iteration (the exact number depending upon the process chosen) thus delaying their opportunity to respond to other events.

Although subjects may take action with respect to no more than one process each iteration, they will be asked to enter a string of proposed actions. The first action in the string of responses represents the actual decision (to continue monitoring or to check one of the nine processes for the occurrence of an abnormal event). The remainder of the string lists the actions subjects would make given that each of the previous actions in the string could not be taken at that time. This sequence will provide information on the extent of the subjects' planned actions.

In this experiment various event occurrence probabilities, advances of process histories on next iteration, and costs for delays in responding to events will be associated with each of the processes. The values of these parameters will be displayed with the process histories to allow subjects to utilize them in making their attention allocation decisions.

The results of this experiment along with those of the first experiment will be used to validate and extend the proposed models of human performance in event detection and attention allocation in a multi-process monitoring situation. The implications of these models to the design and implementation of a computer aiding system will then be considered, perhaps through the use of simulation experiments in which the models are used with a rudimentary computer aid to improve its event detection performance and reduce its tendency to conflict with the human.

IV. References

Afifi, A. A., and Azen, S. P. (1972) Statistical Analysis, New York: Academic Press, 1972.

Carbonell, J. R. (1966) "A Queueing Model of Many-Instrument Visual Sampling," IEEE Transactions on Human Factors in Electronics, Vol. HFE-7, No. 4, pp. 157-164, December 1966.

Carbonell, J. R., Ward, J. L., and Senders, J. W. (1968) "A Queueing Model of Visual Sampling," IEEE Transactions on Man-Machine Systems, Vol. MMS-9, No. 3, pp. 82-87, September 1968.

Curry, R. E., Kleinman, D. L., and Hoffman, W. C. (1975) "A Model for Simultaneous Monitoring and Control," Proceedings of the Eleventh Annual Conference on Manual Control, NASA Ames Research Center, pp. 33-63, May 1975.

Freedy, A., Hall, F. C., Lucaccini, L. F., and Lyman, J. (1971) "A Computer-Based Learning System for Remote Manipulator Control," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-1, No. 4, pp. 356-363, October 1971.

Freedy, A., Weisbrod, R., and Weltman, G. (1973) "Self-Optimization of Task Allocation in Shared Man-Computer Control," Proceedings of the 1973 IEEE Conference on Decision and Control, San Diego, pp. 784-789, December 1973.

Freedy, A., Weltman, G., and Lyman, J. (1972) "Interactive Aspects of a Man/Learning System Control Team," Proceedings of the 1972 IEEE Systems, Man and Cybernetics Society Conference, Washington, pp. 135-140, October 1972.

Gai, E. G., and Curry, R. E. (1975) "Failure Detection by Pilots During Automatic Landing: Models and Experiments," Proceedings of the Eleventh Annual Conference on Manual Control, NASA Ames Research Center, pp. 78-93, May 1975.

Gai, E. G., and Curry, R. E. (1976) "A Model of the Human Observer in Failure Detection Tasks," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-6, No. 2, pp. 85-94, February 1976.

Kleinman, D. L., and Curry, R. E. (1976) "Some New Control Theoretic Models for Human Operator Display Monitoring," Proceedings of the 1976 IEEE Conference on Decision and Control, Clearwater, December 1976.

Kvalseth, T. O. (1975) Quantitative Modelling of the Time-Multiplexing Characteristics of Human Controllers, NSF Report GK-37419, August 1975.

Miller, D. C., and Elkind, J. I. (1967) "The Adaptive Response of the Human Controller to Sudden Changes in Controlled Process Dynamics," IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 3, pp. 218-223, September 1967.

Poulton, E. C. (1974) Experiments on Detection and Classification with the Assistance of a Computer, Summary Report on Sonar 1966-1974, Royal Naval Personnel Research Committee, London, England, AD-A019 314, March 1974.

Rouse, W. B. (1975) "Human Interaction With An Intelligent Computer in Multi-Task Situations," Proceedings of the Eleventh Annual Conference on Manual Control, NASA Ames Research Center, pp. 130-143, May 1975.

Rouse, W. B. (1976a) "A Model of the Human as a Suboptimal Smoother," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-6, No. 5, pp. 337-343, May 1976.

Rouse, W. B. (1976b) "Adaptive Allocation of Decision Making Responsibility Between Supervisor and Computer," in T. B. Sheridan and G. Johannsen, Eds., Monitoring Behavior and Supervisory Control, New York: Plenum, 1976.

Rouse, W. B. (1976c) "Human Computer Interaction in Multi-Task Situations," IEEE Transactions on Systems, Man and Cybernetics, in press.

Rouse, W. B., and Greenstein, J. S. (1976a) "An Experimental Situation for Study of Human Decision Making in Multi-Process Monitoring," Proceedings of the Twelfth Annual Conference on Manual Control, University of Illinois at Urbana-Champaign, May 1976.

Rouse, W. B., and Greenstein, J. S. (1976b) "A Model of Human Decision Making in Multi-Task Situations: Implications for Computer Aiding," Proceedings of the 1976 IEEE Systems, Man and Cybernetics Society Conference, Washington, pp. 425-433, November 1976.

Senders, J. W. (1964) "The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems," IEEE Transactions on Human Factors in Electronics, Vol. HFE-5, No. 1, pp. 2-5, September 1964.

Senders, J. W., and Posner, M. J. M. (1976) "A Queueing Model of Monitoring and Supervisory Behaviour," in T. B. Sheridan and G. Johannsen, Eds., Monitoring Behavior and Supervisory Control, New York: Plenum 1976.

Sheridan, T. B., and Ferrell, W. R. (1974) Man-Machine Systems, Cambridge, Massachusetts: MIT Press, 1974, Chapter 19.

Smallwood, R. D. (1967) "Internal Models and the Human Instrument Monitor," IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No. 3, pp. 181-187, September 1967.

Steeb, R., and Freedy, A. (1976) "Man Machine Interaction in Adaptive Remote Systems," Proceedings of the 1976 IEEE Systems, Man and Cybernetics Society Conference, Washington, pp. 727-731, November 1976.

Tatsuoka, M. M. (1971) Multivariate Analysis, New York: Wiley, 1971.

Weisbrod, R. L., Davis, K. B., and Freedy, A. (1975) "Adaptive Utility Assessment in Dynamic Decision Processes: An Experimental Evaluation of Decision Aiding," Proceedings of the 1975 IEEE Systems, Man and Cybernetics Society Conference, San Francisco, pp. 302-308, September 1975.

Willsky, A. S. (1975) A Survey of Design Methods for Failure Detection in Dynamic Systems, Massachusetts Institute of Technology, Report ESL-P-633, November 1975.

THE PILOT AS A PROBLEM SOLVER

The discussion in the previous three sections of this report have focused on the pilot and computer cooperating in decision making. In this section, we want to discuss the pilot's role as a problem solver in automated and semi-automated systems. First, we will consider the pilot's problem solving tasks in current flight management situations and then, we will discuss how automation may result in the pilot having new and different problem solving tasks.

If one reads the "emergency procedures" sections of aircraft manuals,* one finds that most of the procedures can be generalized as searches of tree-like checklists. Because emergencies allow little time for action, the pilot must memorize the various procedures and then perform them by rote. While he may have some written material available for reference, he can not afford investing too much time in reading. Thus, the pilot must be well-trained in the execution of emergency procedures.

Other problem solving situations for the pilot include course charges due to degradation of aircraft readiness (e.g., a lost engine) or unexpected turbulence. Another example is when conditions change from IFR to VFR (e.g., on approach) and the pilot finds out that the aircraft is not where he thought it was. A further example might be when the pilot unexpectedly finds another aircraft in his vicinity.

In all of the above examples, the pilot is highly trained to respond appropriately. Computer aiding might be of use to the pilot in these

---

* The manuals read were those for the F-4, T-37, and T-38.

situations if it could rapidly supply information necessary to successful solution of the problem. Also, the computer might perform many of the checklist procedures. Further, it could perhaps sort out the numerous warning signals that might be precipitated by the problem solving situation.

It is quite possible that automation will eventually perform the tasks noted in the previous paragraph and it is interesting to consider the implications of such automation. To clarify the issues, let us discuss the pilot's problem solving role in an aircraft where all problem solving is automated. First of all, why would there be a pilot in an automated aircraft? The reason is quite simple. The computer may encounter a task (perhaps, for example, with odds of 1 in 10^6) that it is ill-equipped to deal with (e.g., a hijacking) or the computer may simply fail. Then, the pilot may have to take over manual control or at least solve the particular problem that has stymied the computer.

When the situation requires the pilot to be the problem solver, he is faced with the difficulty of not only dealing with the complexity that existed before automation but also the complexity of the automation itself. Further, if the complexity of automation increases as it becomes more versatile and reliable, then the pilot will increasingly be faced with the following dilemma. As the versatility and reliability of automation increases, the frequency of the pilot having to act as a problem solver decreases while the complexity of the task he must perform increases.

To cope with this dilemma in the design of flight management systems, one will have to accept the fact that the pilot will no longer be experienced in problem solving in the sense that he will not have had numerous replications of practice in solving each possible problem. This conclusion leads to the following question: Should displays, controls, procedures, and computer aiding for the inexperienced* problem solver be different than those for the experienced problem solver?

The first step in our approach to answering this question has been to review the literature. A computer search of Psychological Abstracts yielded citations for 2000-3000 papers on human problem solving published over the past ten years. We chose to review somewhat over 300 recent publications. We have not yet organized this material into reportable form. At this point, we will limit our comments on this literature to one general statement. It appears that displays, etc. for inexperienced problem solvers should be different than those for experienced problem solvers. While the literature reviewed certainly does not provide any design guidelines, it does point out the most important hypotheses to be investigated in an experimental study of pilot problem solving in automated flight management situations.

Our current efforts in this area are aimed at developing an experimental scenario that captures the essence of the problem solving tasks we have discussed above. Our goal is to provide empirical results upon which the design of problem solving displays, controls, procedures, and computer aids can be based. To conclude, in this section we have pointed out what we feel will eventually be a very important problem in flight management. In our next report, we hope to provide more information on the relevant literature as well as a discussion of completed and planned research efforts in this area.

---

*We want to emphasize that the word "inexperienced" is being used to mean someone who may perhaps be well-trained in problem solving procedures but nevertheless has not encountered many of the specific problems that he may be asked to solve.

APPENDIX

Publications based on research supported by NASA Grant NSG-2119

1. W. B. Rouse, Y-Y. Chu, and R. S. Walden, "An Experimental Situation for Study of Pilot Interaction With Automated Airborne Decision Making Systems", Proceedings of the Twelfth Annual Conference on Manual Control, Urbana, IL: University of Illinois, May 1976, NASA TM X-73, 170.

2. W. B. Rouse and J. S. Greenstein, "A Model of Human Decision Making in Multi-Task Situations: Implications for Computer Aiding", Proceedings of the 1976 International Conference on Cybernetics and Society, Washington, November 1976, pp. 425-433.

3. W. B. Rouse, "Human-Computer Interaction in Multi-Task Situations", IEEE Transactions on Systems, Man and Cybernetics, in press.