ARTICLE NO. 78-1589R

Computer-Implemented Grading of Flight Simulator Students

F. L. Comstock*
Singer Company, Binghamton, N. Y.
and
Hugo J. J. Uyttenhove†
State University of New York, Binghamton, N. Y.

This paper presents an investigation into computer monitoring of aircrew performance and subsequent evaluation of that performance. The study presumes that it is possible to develop, for each training task, a general system model of aircrew skill development within the mathematical framework of set theory. Through the judicious selection of parameters to be monitored, and by the utilization of a reasonable amount of computer resources, an evaluation can be made by comparing student performance to a performance optimum. The technique involves the development of a state transition structure difference matrix from the matrix representing the student's performance and the matrix representing the optimum performance.

I. Introduction

SIMULATION is playing an increasingly important role in aircrew training for both military and commercial aviation. Many studies have established that the level of transfer of training to today's high-fidelity simulators is extremely close to a one-to-one ratio in terms of training hours expended.[1]

Moreover, training situations can be created in simulators that can never be duplicated in the real world. An often-cited example is emergency training, for which it is possible to establish simulated conditions offering valuable training in response to crisis situations that would be far too risky to duplicate in actual aircraft. Another example is in combat training. Simulation allows the concoction of a training scenario that just could not be contrived in a military exercise for three reasons: 1) its complexity, 2) the equipment of the foreseen adversary is not available, and 3) the cost.

With the increasing substitution of simulator training for aircraft training there have come significant developments in automated training features, made possible primarily by relatively inexpensive but powerful computers. These automated training features, in the nature of preprogrammed missions, demonstrations, automatic malfunction activation, and the like, have done much to alleviate the workload and increase the effectiveness of the simulator instructor by relieving him of the necessity for constant problem control inputs, thereby allowing him more time for monitoring, guiding, and evaluating student performance. But the burden on simulator instructors remains significant. Work is currently underway to provide the simulator instructor with tools for making the tasks of monitoring, guiding, and evaluating student performance less subjective. This work is frequently referred to as computer-aided instruction (CAI). This paper addresses a portion of that work, specifically that area which is relevant to evaluation.

The problem of computer-implemented performance measurement has been addressed by Connelly et al.[2-5] Their comprehensive investigation includes three computational models. The first measures trends in performance by means of recording state transfers; the second compares performance with an absolute performance reference; and the third treats invariant relationships among separate variables in the performance data. Using statistical methods, they analyzed their data to determine the accuracy and relevancy of their score predictions for the three models with excellent success. Nevertheless, their techniques require the availability of large amounts of computational resources.

It is our belief that a method of grading performance, which handles the unpredictable maneuvers and tasks as well as the predictable at a reasonable cost in terms of computer resources, is essential to performance evaluation in flight simulators. To accomplish this, we have turned to the field of general systems theory to find the proper methodology and tools with which to investigate the evaluation process. In particular, we have applied a General Systems Problem Solver[6] (GSPS) which is assumption free in its nature. The mathematical tools we used, especially with respect to placing the transitions from one performance condition to another in matrix form, are similar to the first model of the aforementioned researchers. However, our matrices are uniquely developed and utilized. The GSPS concepts involved are extensive, but crucial to our approach; their basic outlines are presented in the following section.

II. Relevant Concepts

General Systems Problem Solver

The problem-solving approach by the GSPS is based on a framework of taxonomies of systems, problems, methodological tools which provide solutions to the problems, and some systems modeling aspects.

The most fundamental classification is that of epistemological levels of systems, as suggested by Klir.[7-9] The basic hierarchical categories are:

1) Source systems (dataless sytems) which define the object of investigation in terms of attributes and variables, together with their possible appearances and states, respectively, and an interface between the attributes and variables.

2) Data systems consist of measured, observed, or given data for the variables, usually appropriate parameters of some kind, such as time, space, population, etc.

Presented as Paper 78-1589 at the AIAA Flight Simulation and Technologies Conference, Arlington, Texas, Sept. 18-20, 1978; submitted Nov. 27, 1978; revision received April 3, 1979.

Index categories: Simulation; Cabin Environment, Crew Training, and Life Support.

*Staff Engineer, Training Systems Engineering, Link Division.

†Program Analyst, Dept. of Systems Science, School of Advanced Technology; currently on a Research Grant with Inter Universitaire Interfaculteit, The Netherlands.

3) Behavioral systems or generative systems in the sense that the data systems can be predicted or generated by means of parameter invariant procedures.

4) Structure systems defined in terms of a set of elements and some relation between them, such that behavioral or lower-level systems are fully described.

5) Metasystems at various levels which are defined in terms of known systems corresponding to lower epistemological levels, together with a parameter invariant procedure which describes changes from one system to another.

At each level, the classification is extended by inclusion of various methodological distinctions. For further classifications, see Cavallo and Klir[6] and Uyttenhove.[10]

The taxonomy of problems consists of problem types and particular problems defined in terms of system types, particular systems, and requirements. Solutions for particular problems are provided, if possible, by the methodological tools. The latter are an assemblage of interactive computer programs which guide the investigator or modeler through the final stages of the GSPS.

Thus, we are dealing with systems and the identification of them at various levels of the hierarchy of epistemological levels.

Regardless of the level at which the system is defined, the variables involved in the system may be classified into input and output variables. This classification means that states of the input variables are viewed as circumstances under which states of output variables are conditional: "If the input variables are in state x, then...." Input variables are, therefore, not the subject of inquiry, but are viewed as being determined by some agent which is not part of the system under consideration. Such an agent is referred to as the environment.

Systems whose variables are classified into input and output variables are called directed systems; those for which no such classification is given are called neutral systems. This dichotomy of systems holds for each of the epistemological levels.

The methodological criteria regarding the kinds of variables and relations are also part of the systems at each of the epistemological levels. We have already noted the distinction between input and output variables, while we also consider dichotomies such as variables being well defined or fuzzy, discrete or continuous, of nominal or ordinal scale, linearly or nonlinearly ordered parameters, etc. Relations can be classified as being deterministic or probabilistic, memoryless or memory dependent, linear or nonlinear, etc.

These base categories enable us to describe the mode of operation of GSPS in a more direct way. We must realize, however, that the computer-implemented GSPS, i.e., its interaction facilities and the tools, is constantly subject to changes in the sense that it (GSPS) learns from experience and as such stores information about problem types and particular problems for future reference. Furthermore, GSPS broadens its horizon by its expansion as an effort to solve more and larger problems.

In principle, the operation of the implemented GSPS consists of accepting inputs (i.e., the problem statement) and generating the correct sequence of tools and their output solution. Specifically, the output can be either a particular system of the type demanded by the problem, or a relation between the initial and terminal systems. With regard to the input into GSPS, some interactive procedure should assist the investigator in identifying the entry to the GSPS. This means that the language and interaction should help to formulate the problem in an understandable general systems expression. The admissibility of the problem is largely dictated by the methodological tools. These, together with the specific sets of epistemological and methodological criteria, are coded on computer files. An interactive program activates the communication between the investigator and the user. The following sections of the algorithm are described from the standpoint of the GSPS.

1) Find out what the investigator defines as the initial system type and identify a particular system of that type.

2) Determine the requirement types and the set of particular requirements from answers by the investigators.

3) Conclude, if possible, what the terminal system type is and, if the problem so requires, identify a particular system of that type.

4) Determine whether the problem is admissible and, if so, which tools will provide the solutions.

5) If no tool can be assigned or if the problem is not admissible, determine if information relevant to the problem is available and allow for respecifications of steps 2-4, or the GSPS can offer no solution.

That part of the GSPS which is relevant to this paper is the description of systems at levels 1, 2, and 4 of the hierarchy. In the following section, we discuss some relevant formal concepts which are the backbone of the approach.

Formal Notions

Let $A = \{a_i | i \epsilon I_n\}$ be a set of basic attributes chosen by the investigator to represent an object of interest for some specific purpose; $I_n = \{1,2...,n\}$ is an index set determined by the number n of basic attributes. Let A_i denote the set of potential appearances of the basic attribute a_i.

Let $B = \{b_j | j \epsilon I_m\}$ be a set of supporting attributes (time, space, etc.) chosen by the investigator and let B_j stand for the set of potential appearances of supporting attribute b_j. Then, the neutral object system in the simplest form is defined by the pair

$$\Omega_N = [\{(a_i, A_i) | i \epsilon I_n\}; \{(b_j, B_m) | j \epsilon I_m\}]$$

Any useful relation (e.g., ordering) which can be recognized in the sets of appearances of the attributes should be added to the definition of the object system. Moreover, if there exist dependencies among the basic attributes which are solely due to observation or measurement procedures (e.g., one attribute represents an arithmetic average of others), then these dependents should also be included in the definition.

Let v_i and V_i $(i \epsilon I_n)$ denote basic variables and sets of states of the variables, respectively, and let supporting variables and their sets of states be denoted by w_j and W_j $(j \epsilon I_m)$, respectively. Note that $V = \{v_i | i \epsilon I_n\}$ and $W = \{w_j | j \epsilon I_m\}$. A neutral image system Γ_N, compatible with the neutral object system Ω_N, is then defined as the pair

$$\Gamma_N = [\{(v_i, V_i) | i \epsilon I_n\}; \{(w_j, W_j) | j \epsilon I_m\}]$$

This definition may be supplemented by some recognized relations in the sets of stores and/or dependencies among the basic variables. The set $\underset{j \epsilon I_m}{\mathrm{X}} W_j$ is referred to as the parameter space ω.

Assume now that a resolution level is introduced for each attribute of the object system which characterizes the meaning of data to be collected for the attribute. This can be accomplished by defining a partition π_i (A_i) on each set A_i $(i \epsilon I_n)$ and a partition $\pi_j(B_j)$ on each set B_j $(j \epsilon I_m)$. The form each of these partitions takes depends primarily on the knowledge of the object of investigation, the purpose of the investigation, and the measuring instruments, as well as computing facilities which are available.

In order to get a meaningful basis for data gathering and data interpretation, a correspondence between the entities involved in the object and image systems must be introduced. This is accomplished by: 1) a one-to-one correspondence f_a: $A \rightarrow V$; 2) a one-to-one correspondence f_b: $B \rightarrow W$; 3) a family of one-to-one correspondences $G = \{g_i: \pi_i(A_i) \rightarrow V_k | i,k \epsilon I_n; v_k = f_a(a_i)\}$; 4) a family of one-to-one correspondences $H = \{h_j: \pi_j(B_j) \rightarrow W_l | j,l \epsilon I_m; w_l = f_b(b_j)\}$. The collection of an object system, an image system, and a correspondence between them expressed in terms of the one-to-one correspondences 1-4 form a complete frame for data gathering and interpretation referred to as the neutral source system.

Each dataless system (image or source system) implicitly contains all possible trajectories of states of basic variables in the parameter space, i.e., all possible functions from

$$\underset{j\epsilon I_m}{\mathrm{X}} W_j \text{ to } \underset{i\epsilon I_n}{\mathrm{X}} V_i$$

A meaningful restriction to one of the functions (for example, function δ), which in the modeling problem is determined by data gathering, constitutes data regarding the variables. When the neutral dataless systems, say $^0\mathrm{S}$ is augmented with δ, we obtain a neutral data system (or a neutral system at epistemological level 1). Let $^1\mathrm{S}$ stand for the neutral data systems, then,

$$^1\mathrm{S} = (^0\mathrm{S}, \delta)$$

(Systems $^i\mathrm{S}$ implicitly stand for neutral systems unless specifically subscripted by N or D (neutral or directed) for reasons of differentiation.)

Let a set of variables $s_k\,(k\epsilon I_q)$, referred to as sampling variables, be introduced by the equation

$$s_{k,w} = v_{\lambda_r}(w)$$

where $s_{k,w}$ stands for states of sampling variable s_k at point w in the parameter space, and λ_r denotes a parameter-invariant translation rule, which for any given point w in the parameter space determines one or several other points in the parameter space.

For instance, when the parameter space is totally ordered (as in the case of time parameter) and represented by the set of positive integers, each translation rule can be described by a simple equation

$$\lambda_r(w) = w + a$$

where a is an integer. Sampling variables are then defined by

$$s_{k,w} = v_{i,w+a}$$

When the parameter space is partially ordered, $\lambda_r(w)$ may stand, for example, for points in the parameter space which are predecessors (or successors) of w with a particular distance from w.

Let Λ denote the set of all translation rules under consideration and let the relation

$$M \subset V \times \Lambda$$

specify which translation rules are applied to which variables (including possible internal variables). Set M is called a mask.

Given a dataless system and a mask, a set of sampling variables is uniquely defined. A relation $R_I \subset S$ defined on

$$S = \underset{k\epsilon I_q}{\mathrm{X}} S_k$$

then can be introduced. Elements of R_I are q-tuples of states of the sampling variables defined by the mask. They are called data samples. When probabilities $p(c)$ are given with which samples $c\epsilon S$ appear, we obtain set

$$\mathrm{B}(bb) = \{(c,p(c)) \mid c\epsilon R_I, 0 \le p(c) \le 1, \sum_c p(c) = 1\}$$

which is referred to as basic behavior of the behavioral system under consideration.

To employ the basic behavior for generating data for the given primitive system, some order of states $w\epsilon\omega$ must be chosen in which the data are generated. The order must be compatible with the natural order of the parameter space ω. If the parameter space has no natural order, it may be ordered artificially in some suitable way. Given an order of states in ω (linear or partial), we let $w \le w'$ denote that either state $w\epsilon\omega$ precedes state $w'\epsilon\omega$ or $w = w'$.

Once a generative order in set ω is decided, a relation

$$R_2 \subset R_I \times R_I$$

whose elements are pairs (c,c') of successive samples with respect to the generative order, can meaningfully be defined. When probabilities $p(c,c')$ are given with which pairs (c,c') appear, we obtain a set

$$\mathrm{B}(bst) = \{[(c,c'), p(c,c')] \mid (c,c')\epsilon R_2,$$

$$0 \le p(c,c') \le 1, \sum_{c,c'} p(c,c') = 1\}$$

This set is called the basic state-transition (ST) relation.

The neutral behavioral system $^2\mathrm{S}$ is defined as a triple

$$^2\mathrm{S} = (^0\mathrm{S}; M; \mathrm{B})$$

where $^0\mathrm{S}$ is a neutral dataless system with ordered set ω, M denotes a mask, and B stands for an element taken from the set $\{\mathrm{B}(bb), \mathrm{B}(bst)\}$.

Although the behavioral system does not explicitly contain any data, it contains a relation through which data can be generated. Hence, it consists of a source system, a generative relation invariant with respect to the state set ω of supporting variables, and, indirectly, a set of data systems which can be generated through the relation. For the formal concepts regarding structure systems $^3\mathrm{S}$, we refer to Klir and Uyttenhove.[9]

When behavior or structure systems offer no solution to the problem investigated, one of the reasons may be that the parameter invariance does not hold. In other words, if time is a supporting parameter, then the systems may turn out to be time dependent. To determine the points in the data system where this dependency becomes evident, an algorithm was developed by Uyttenhove.[10,11] For our application purposes, the basic feature is based on the probability distribution for the samples c on which we calculate the uncertainty as

$$H = -\Sigma p(c) \cdot \log_2 p(c)$$

The samples are collected for larger and larger segments ($\hat{s}$) until the measurement for two successive segments shows a percentage change larger than an acceptable threshold, i.e., $\Delta H > \tau$. This process is carried out for different segment increment sizes $\hat{s}$ until a stable pattern emerges. For the nontrivial case of general parameter variance and larger masks, we refer to previous work.[10] Systems which are composed of a set of lower-level systems and for which the procedure is known by which they "connect," are called metasystems. The metasystems is defined as a triplet

$$^4\mathrm{S} = (T, S, P)$$

where T is the parameter space, S the set of lower-level systems and P the procedure.

System Identification Tools

The GSPS contains several tools that will aid us in the identification of the proper systems. Identifying the behavioral system in state transition mode is a trivial operation. Identifying the metasystem is accomplished by means of an algorithm META. During a full-length application of the GSPS, the investigator would be directed to these tools via an interactive program on a terminal time-

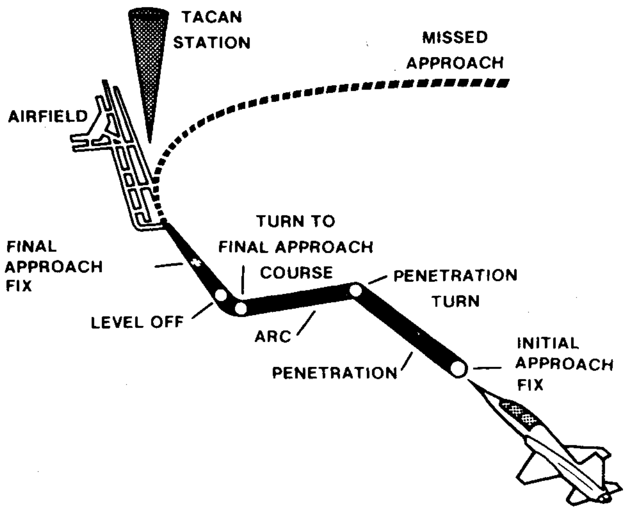

Fig. 1 Typical TACAN approach.

sharing device. Since it is our intent here to briefly state the GSPS concepts relevant to our study, we refer to the previous references for full details of the methodology.

III. The Problem and Its Solution

Problem Statement

The example chosen to illustrate the concepts presented in this paper is a relatively simple one. Extending the concepts to more complex and less predictable maneuvers requires further work, but the basic framework exists here.

The data used are for a typical tactical air navigation (TACAN) approach (see Fig. 1) as performed in a jet trainer. The attributes selected do not necessarily reflect all parameters that comprise the descriptors of a typical instrument approach, but they do define adequately the essentials of an instrument approach. The intent is to describe a defined flight path relative to a fixed ground reference and monitor how well a student can control the simulator so that it adheres to that flight path.

The parameters of altitude, TACAN radial and distance measuring equipment (DME) reading taken together totally describe the desired flight path for all points from the initial approach fix to the missed approach point (MAP). The airspeed parameter is a regulatory one imposed by controlling agencies and/or aircraft performance characteristics. It was desired for purposes of this example to limit the monitored attributes to a small set; however, the system is capable of handling more attributes, such as those describing aircraft attitude and procedure accomplishment.

The technique for student evaluation is based on a computer comparison of the ST behavior, B (*bst*), for the aircraft under the control of the student against the optimum ST behavior established by the published approach path and procedures. The optimum values for a maneuver, taken at specific points and reduced to an ST matrix, are stored in the simulator computer. The parametric values for the simulator under student control are recorded as the student's flight progresses and these data are reduced similarly. The comparison of the stored optimum ST matrix is then made against the student's ST matrix. The student's grade is a matter of establishing his deviations from the optimum ST structure probabilities and outputting a standardized score based on a large data base that can easily be developed over a period of time.

Using the referenced data, an activity matrix was developed for a typical TACAN approach by a jet training airplane.

Data System

The generated data matrix used as the standard in this study follows standard instrument procedures and covers the period from arrival at the initial approach fix (IAF) through the missed approach. The appearances of the variables, airspeed, altitude, TACAN radial, and TACAN distance or DME were described, respectively, in the common aviation units of knots, feet above sea level, degrees relative to magnetic north, and nautical miles. Values were established for points approximately 5 s apart assuming a no-wind condition. A sample of the generated data array is given in Table 1. It should be clearly understood that in this study the flight path generated and used as the standard against which student performance is to be measured was strictly hypothetical. The variable values making up the activity array were estimates based on the writers' experience and judgment and do not represent actual recorded values. In implementation, the variable values used as a standard would be generated and recorded during actual aircraft or simulator flights performed by one or more highly qualified USAF flight instructors.

Subsequent to the data collection, we carried out the mapping from appearances of the basic attributes to the states of the basic variables, according to set ranges of values (see Table 2). These ranges were not selected arbitrarily, but were carefully chosen to create windows of acceptable performance at each of the points in the approach. For example, with respect to airspeed, one range covers the penetration airspeed with a tolerance of -10 (no upper limit), one range covers the band of recommended final approach level-off airspeeds (220-249), while a third covers the final approach airspeed with a tolerance of -5 knots and $+14$ knots to account for airspeed adjustments because of fuel. The other airspeed ranges cover the transitions from one approach phase to another. The tolerance bands can be tightened or loosened by adjustments in the ranges according to the proficiency level. The final approach airspeed range can also be more accurately modeled by requiring the computer to calculate the proper approach airspeed based on fuel remaining, and applying the tolerance band to that figure.

Metasystem Identification

The data system served as input to the META algorithm in order to determine points of change in the uncertainty measure. Although we only explored the computation for masks of depth 1, we likewise carried out the identification procedure for masks of depth 2. Table 3 summarizes the results.

The selection of $\hat{s}$ was based on the fact that $4 < |T|/\hat{s} \leq 13$ was desirable from the point of view that 170 observations were available. Because of diversity in results and because points are difficult to identify uniquely, we average the results. These results conform with the segments of the total approach. System 1, for data from point 1 to 70, represents the penetration descent; system 2, for data from point 71 to

Table 1 Sample data array

	Time interval						
	0	1	2	3	4	5	6
Airspeed	250	250	250	265	280	280	280
Altitude	23,000	23,000	23,000	22,700	22,100	21,257	21,414
Radial	102	087	082	082	082	082	082
DME	34.6	34.3	34.0	33.7	33.4	33.0	32.6

Table 2 Mapping of appearances to states

Attributes	Variables Order	MNEM	1	2	3	4	5	6	7	8	Support
Airspeed	1	S	150-169	170-219	220-249	250-269	>269	<150			Time is the supporting variable
Altitude	2	A	1700-1839	1840-2399	2400-2699	2700-4899	4900-6899	6900-22,800	>22,800	<1700	
Radial	3	R	0.0-71.9	72.0-92.9	93.0-112.9	113.0-121.9	122.0-360.0				
DME	4	D	0.00-2.59	2.60-13.59	13.60-20.59	>20.59					

Table 3 META system identification results

Depth of mark	Tolerance, τ	Increment segment size $\hat{s}$	Points indicating uncertainty change $>\tau$, each point denotes beginning new system: System 1	System 2	System 3
1	10%	13	1	79	
		17	1	52	120
		21	1	64	127
		29	1	88	...
		42	1	...	127
2	10%	13	1	79	117
		17	1	69	120
		21	1	64	127
		29	1	...	117
		42	1	...	...
	Average		1	71	123

122, represents the flight along the arc and interception of the final approach course; system 3 is akin to data collected from the final and missed approach.

Since the metasystem identification was carried out successfully, we could continue investigating each of the smaller data systems on an individual basis.

Behavioral Systems (State Transition)

For each data system, an ST behavior [B (*bst*)] was obtained in matrix form. The matrix elements are probabilities associated with the plane moving to a particular state, given that it was in a certain state just before. By inspection of Fig. 2 (a, b, and c), we notice that transitions are either to the same or one other state. This was expected since the airplane moved along a flight path that never intersected an earlier portion of the flight path nor, since the flight path was optimized, did the airplane ever leave the desired flight path (state) and then correct back again. In other words, a state was never entered more than once in the optimum approach.

Conclusion

With the aid of the ST matrices, we will now be able to carry out the evaluation of student pilots. Before operating on this aspect of the study, let us first return to the ST matrix and see what we can expect. First of all, the airplane is ruled by the laws of physics and, therefore, the variables representing its location and movement cannot jump abruptly in value, but must represent a gradual transition. There are no right-angle turns or discontinuities in airspeed, altitude, or ground track. It is a continuum and the transition of each state, either to itself or to some adjacent state, is normal. It might, therefore, be argued that the whole system can represent the airplane's flight path just as well as the smaller systems. However, it must be remembered that one of the primary purposes of a grading system is to provide a learning tool. A good grading system should do more than rank students. It should also point out, to student and instructor alike, the student's weak and strong areas. When applied to flight training, this means that grades should identify the areas in which a student should be required to attain a higher level of skill development, an obvious need in a field in which substandard skill levels can be, and are, all too frequently, fatal. To provide the detailed identification of weaknesses that is demanded, more than one overall grade is required. This leads quite naturally to task segmentation with grading of each segment and hence to subsystems of the overall system that represents the flight maneuver. This segmentation is part of the considerations of implementing the obtained results for the GSPS tools.

IV. Implementation Considerations

Hypothetical Implications

Now that a model for optimum performance exists, student evaluation is a matter of comparing his behavior with that of the model. The student's performance is recorded in the computer in terms of the data system, and the ST matrix is calculated. By taking the difference between the optimum ST matrix and the student's ST matrix and summing the absolute difference values of the individual matrix entities, a raw grade within the 0-2 range is obtained.

This range results because the probabilities of each ST matrix sum to one. Thus, a student who follows the optimum flight profile will have a grade of zero (no differences), and a student who is constantly out of tolerance and therefore is constantly transitioning to an incorrect state will have a grade of two (no commonality with the optimum flight path). The criteria on which every student's evaluation is made are: 1) does he attain the proper states: 2) does he remain in each proper state through the appropriate number of recorded points; and 3) does he transit to the next proper state when leaving a given present state.

The appearance of the correct states in the student's ST matrix establishes that he attained the correct states. The probability differences between the student's matrix and the optimum matrix, for those states which represent a transition to themselves, determine if the student remained in a given state through the proper number of recorded points. The probability differences between the student's matrix and the optimum matrix for the state transitions to new states

(a) NEXT STATE

PRESENT STATE	s8	s7	s6	s3	s2	s5	s4
s8	.22	.00	.00	.00	.00	.00	.00
s7	.01	.02	.00	.00	.00	.00	.00
s6	.00	.01	.07	.00	.00	.00	.00
s3	.00	.00	.00	.00	.00	.00	.01
s2	.00	.00	.00	.01	.02	.00	.00
s1	.00	.00	.00	.00	.01	.00	.00
s5	.00	.00	.01	.00	.00	.14	.00
s4	.00	.00	.00	.00	.00	.01	.46

(b) NEXT STATE

PRESENT STATE	s14	s13	s12	s11	s10	s9	s8
s14	.46	.00	.00	.00	.00	.00	.00
s13	.02	.04	.00	.00	.00	.00	.00
s12	.00	.02	.18	.00	.00	.00	.00
s11	.00	.00	.02	.02	.00	.00	.00
s10	.00	.00	.00	.02	.02	.00	.00
s9	.00	.00	.00	.00	.00	.04	.02
s8	.00	.00	.00	.00	.02	.00	.14

(c) NEXT STATE

PRESENT STATE	s17	s16	s15	s18	s19	s20	s21	s22
s17	.09	.00	.00	.02	.00	.00	.00	.00
s16	.02	.04	.00	.00	.00	.00	.00	.00
s15	.00	.02	.11	.00	.00	.00	.00	.00
s14	.00	.00	.02	.00	.00	.00	.00	.00
s18	.00	.00	.00	.09	.02	.00	.00	.00
s19	.00	.00	.00	.00	.00	.02	.00	.00
s20	.00	.00	.00	.00	.00	.04	.02	.00
s21	.00	.00	.00	.00	.00	.00	.07	.02
s22	.00	.00	.00	.00	.00	.00	.00	.40

(d) STATE

STATE	
4734	s1
4724	s2
4624	s3
5624	s4
5623	s5
4623	s6
4523	s7
3533	s8
3543	s9
3443	s10
2443	s11
2442	s12
1442	s13
1342	s14
1242	s15
1142	s16
1141	s17
2241	s18
3251	s19
3351	s20
3451	s21
3452	s22

Fig. 2 State transition matrices for system 1 (a), system 2 (b), system 3 (c) and the state representation in terms of values of the variables (d).

(a) NEXT STATE

PRESENT STATE	s1	s2	s3	s4
s1	0	0.2	0	0
s2	0	0.6	0.2	0
s3	0	0	0	0
s4	0	0	0	0

(b) NEXT STATE

PRESENT STATE	s1	s2	s3	s4
s1	0	0.2	0	0
s2	0	0.2	0.2	0
s3	0	0	0	0.2
s4	0	0	0	0.2

(c)

STATE	
4734	s1
4724	s2
4624	s3
5624	s4

(d)

STATE	OPTIMUM PROBABILITY	STUDENT PROBABILITY	ABSOLUTE PROBABILITY DIFFERENCE
s1-s2	0.2	0.2	0
s2-s2	0.2	0.6	0.4
s2-s3	0.2	0.2	0
s3-s4	0.2	0	0.2
s4-s4	0.2	0	0.2
Total absolute probability differ.			0.8

Fig. 3 Sample ST matrices for student (a), model (b), state representation in terms of the values of the variables (c), and error value calculation (d).

determine if the student transits to the next proper state when leaving a given state. This information can be output in a hard copy format that will identify the point in the maneuver at which the student strayed out of tolerance and the specific parameters that the student was negligent in controlling. This in itself will provide a valuable instructional tool for post-mission debriefing of the student. Not only is the need for reliance on the instructor's memory obviated, but the error data are specific and objective. The hard copy output of this information can be expressed in the vernacular of aircrews for ease of reading so that it can be readily interpreted by instructor and student alike.

Application

For illustrative purposes, we assume a hypothetical behavioral system (i.e., ST matrix) as a given optimal model. We further assume an ST matrix which resulted from the student's interaction in the aircraft simulator. Both ST matrices are shown in Fig. 3. The high probability associated with the student keeping the aircraft in state s2 reflects his delay in descent and acceleration beyond the initial approach fix. Due to this error, the transition to states s3 and s4 have been delayed.

In order to numerically evaluate this error, the absolute-difference values of corresponding matrix elements (probabilities) are calculated for each matrix. This results in an error value of 0.8, which can now be presented to the instructor with optional comments in hard copy format.

If the students are to be ranked by this evaluation scheme, some weighting should be applied to the raw scores. The individual probability differences make no distinction between a student whose incorrect system states differ from the optimum states by only one variable state, and a student whose incorrect system states differ from the optimum state by two or more variable states. This shortcoming can be overcome by multiplying the probability differences for those unique state transitions that appear only in the student's matrix by the number of variable states that differ from the optimum flight path state. This would have the effect of penalizing each student who deviated from the optimum flight path in accordance with the number of parameters he allowed to stray out of tolerance.

If the students are to be graded relative to other students for overall mission performance, further weighting applied to maneuver phases and some form of standardized scoring should be applied. In our example, system 3 represents the most critical phase of the instrument approach, because the aircraft is nearing the runway in preparation for landing and, therefore, altitudes, airspeeds, and adherence to the course line are more critical than in the preceding phases. For this reason, the raw score for performance in system 3 should be weighted more heavily than performance in systems 1 and 2; however, the weighting factor to be applied to each phase of a maneuver is a subjective decision best left to those actively engaged in aircrew instruction.

The conversion of raw scores to a form of standardized scores is a matter of obtaining a sample of sufficient size to produce a statistically reliable base for the scoring curve. Obtaining a sample of valid subjects large enough for reliable results will require that the simulator be in use for training for an appreciable period. For this reason, the development of standardized scores based on the raw scores computed by this evaluation scheme also will be left to the using agencies.

It must be admitted that our illustrative example is based on a highly structured maneuver for which the desired behavior is totally defined. Such an example was purposely chosen to clearly demonstrate the mechanics of this method of evaluation. However, one might logically question how GSPS tools can be applied to evaluate students performing less predictable flight tasks such as those in air combat maneuvering. The response to such a question is that as long as the proper variables and their states can be identified and the problem admissible, GSPS can attempt to complete the task. For instance, numerous studies of the problems of air combat maneuver training have been made[12-14] and, based on these studies and other trainers such as Link's Simulator for Air-to-Air Combat (SAAC) have been designed. The highly successful SAAC employs a program to "fly" a computer-generated target against students flying the fighter cockpits in a simulated environment. This program extrapolates the near-term flight path of the student's simulated aircraft and, based on that extrapolation, selects the most advantageous countermaneuver for the generated target. It would be a small step for the computer to go through an identical decision process to establish the student's optimum next maneuver which is, in essence, the optimum next-system state.

The methodological tools of GSPS impose no constraints on the real-time development of the optimum ST matrix. In other words, at each point in the training exercise, the computer can determine and record the student's optimum next state, as well as record the student's actual present state. These recordings are sufficient to establish the ST matrices on which this method of evaluation is based.

V. Conclusions

General Systems Problem Solving tools show promise as methods for computerized grading the performance of simulator students. Its advantages are:

1) It unburdens the instructor by relieving him of the necessity for maintaining a record of student activity.

2) It is objective in that it is based on the attainment of specific goals.

3) It can provide a valuable postmission debriefing tool in the form of a hard copy record of student performance.

4) It is based on relatively simple concepts.

5) It can be applied to complex training tasks.

Nevertheless, the implementation of student performance evaluation requires a thorough analysis of each maneuver or task, careful selection of the attributes and their appearances, and judicious mapping of attributes and their appearances to variables and their states. Flight parameters and procedures must have specific values and sequences that fit the time, location, and conditions under which the maneuver or task is performed. This evaluation technique is no exception. In order that it be applied, each maneuver must be thoroughly analyzed. This can be an arduous job where something as complex as an air combat situation is to be modeled. However, the first one is always the most difficult, and once the initial maneuver or task in any group is completed, the others come more easily, be they TACAN approaches or air combat maneuvers.

References

[1] Diehl, A.E. and Ryan, L.E., "Computer Simulator Substitution Practices in Flight Training," TAEG Rept. No. 43, 1977.

[2] Connelly, E., Schuler, A., and Knoop, P., "Study of Adaptive Mathematical Models for Deriving Automated Pilot Performance Measurement Techniques," AFHRL-TR-69-7, Air Force Human Resources Lab. (AFHRL), Wright-Patterson AFB, Ohio (2 vols.) AD704597 and AD704115, 1969.

[3] Connelly, E., Schuler, A., Bourne, F., and Knoop, P., "Application of Adaptive Mathematical Models to a T-37 Pilot Performance Measurement Problem," AFHRL-TR-70-45, Air Force Human Resources Lab., Wright Patterson AFB, Ohio, AD726632, 1971.

[4] Connelly, E., Bourne, F., Loental, D., and Knoop, P., "Computer Aided Techniques for Providing Operator Performance Measures," AFHRL-TR-74-87, Air Force Human Resources Lab., Wright Patterson AFB, Ohio, AD014330, 1974.

[5] Connelly, E., Knoop, P., Bourne, F., and Loental, D., "Computer Aided Generation of Performance Measures for Man-Machine Systems," *Proceedings of the Human Factors Society Conference*, Huntsville, Ala., 1974.

[6] Cavallo, R.E. and Klir, G.J., "A Conceptual Foundation for System Problem Solving," *International Journal of Systems Sciences*, Vol. 9, Feb. 1978, pp. 219-236.

[7] Klir, G.J., *An Approach to General Systems Theory,* Van Nostrand, New York, 1969.

[8] Klir, G.J., "On the Representation of Activity Arrays," *International Journal of General Systems*, Vol. 2, No. 3, 1975.

[9] Klir, G.J. and Uyttenhove, H.J., "Computerized Methodology for Structure Modelling," *Annals of Systems Research*, Vol. 5, 1976.

[10] Uyttenhove, H.J., "Metasystem Identification," *Applied General Systems Research*, G. J. Klir (ed.), Plenum Press, New York, 1978, pp. 147-160.

[11] Uyttenhove, H.J., "Computer-Aided Systems Modelling: An Assemblage of Methodological Tools for Systems Problem Solving," Doctoral Dissertation, State University of New York, 1978.

[12] Campbell, T.K. and Hartsook, L.B., "A New Approach to Onboard Real-Time Optimum Computations for Aerial Combat Games," AIAA Paper 71-797, presented at the 3rd Aircraft Design and Operations Meeting, Seattle, Wash., 1971.

[13] Lee, V.A. and Wenham, R.J., "Air Combat Simulation," AIAA Aircraft Design for 1980 Operations Meeting, 1968.

[14] Leatham, A.L. and Lynch, U.H.D., "Two Numerical Methods to Solve Realistic Air to Air Combat Differential Games," AIAA Paper 74-22, presented at the 12th Aerospace Sciences Meeting, Washington, D.C., 1974.