Integrated Modeling of Cognition and the Information Environment

Final Report for Project:

Submitted to:

NASA Ames Research Center
ATTN: David Foyle, Ph.D.
Mailstop 262-4
Systems Safety Research Branch
Building 262, Room 269
Moffett Field, CA 94035-1000

January 2002

Submitted by:

Michael D. Byrne, Ph.D.
Department of Psychology
Rice University, MS-25
Houston, TX 77005-1892
Email: byrne@rice.edu
Voice: 713-348-3770

Alex Kirlik, Ph.D.
Aviation Human Factors Division
University of Illinois, One Airport Road
Savoy, IL 61874
Email: kirlik@uiuc.edu
Voice: 203-799-0443

CONTENTS

Integrated Modeling of Cognition and the Information Environment

Final Report

1. Executive Summary

This report describes the progress made, findings, and lessons learned in modeling the T-NASA2 scenario within the ACT-R/PM cognitive architecture. The report begins with a brief overview of the motivation for integrated modeling of cognition and the information environment in the domain of aviation safety. The ACT-R/PM modeling framework is then introduced, along with a description of the ecological, rational and environmental analysis methods needed to put flesh on the bones of this cognitive architecture. The application to the T-NASA2 scenario is described next, first in terms of overall flow of control (goal selection), and at more detail in terms of the goal decomposition around which ACT-R/PM productions were constructed (e.g., look for incursion, maintain speed).

Analysis and modeling revealed that an accurate representation of turn-related decision strategies was crucial to the success of the model, in terms of replicating human performance and error in the T-NASA2 scenario. As a result, a study was performed with an SME, a working airline pilot, to assess any high level (i.e., airport neutral) knowledge pilots might use in deciding what turns to make during taxi operations. The results of this rational analysis study, which was based on a sample of 284 candidate taxi routes over eight major U.S. airports, suggested that taxiway geometry and airport layout was much more regularly structured than we had originally thought. This finding suggested that it is likely that pilots make turn-related decisions using a combination of specific information gained in real time from perception of airport layout, and from longer term and more general knowledge of regularities in the likely paths of travel between touchdown and gate. In the ACT-R/PM implementation, this was achieved by providing the model with multiple, redundant methods for making turn-related decisions, organized in a cost-benefit (rational analysis) hierarchy. The model attempts to make the most accurate turn-related decision possible (benefit) given the constraints of time available (cost). This Hierarchical Heuristic (HH) strategy for turn-related decision making appeared to produce model behavior consistent with the pattern of errors observed in the T-NASA2 scenario.

The report also describes the ecological and task analyses that were required in order to situate the cognitive model in a plausible dynamic context of temporal constraints and opportunities. The behavior of the HH strategy described above, in terms of actual turn decisions made, was highly dependent on the time available for decision making. The fact that the decision maker being modeled was situated in a 60,000 Kg vehicle with engines capable of providing 20,000 lbs of thrust, attempting to negotiate taxiway turns of 25 m radius or less, was obviously not incidental to achieving a reasonable dynamic estimate of actual decision horizons. As a result, ACT-R/PM cognitive modeling had to be supplemented with at least cursory modeling of the taxiing dynamics of the simulated T-NASA2 aircraft in order to achieve a reasonable representation of the interactive, pilot-aircraft system for actual model runs.

The next section of the report provides a discussion of findings, lessons learned, and recommendations for future research. These fall into two broad categories. First, lessons learned related to technical issues in computational, cognitive modeling are presented. Here, we conclude that significant challenges still exist in bringing scientific modeling approaches, such as ACT-R/PM, to bear on applied human-machine contexts. First, advances appear to be needed to expand the range and representation of inputs and outputs these models use for communicating with the external world, i.e., the world beyond the desktop computer. Second, advances are needed to more readily support the coupling of computational cognitive models with sophisticated, computationally autonomous, models of task-environmental dynamics (e.g., the flight simulator used in T-NASA2 research). A higher degree of parallelism than is currently available appears to be required.

Finally, we describe findings and lessons learned regarding our interpretation of the T-NASA2 data. Despite the modeling challenges outlined above, we do believe that our interpretation of human errors in terms of the HH strategy, coupled with the limited visibility and time constrained nature of the task environment, is a reasonable one, and should be further pursued. Given this modeling demonstration, we hope that this research will convince more members of the human factors and aviation safety community to investigate human performance issues with the benefit of emerging developments in computational modeling.

2. Introduction and Motivation

2.1 Introduction

Aviation incident and accident investigations often find both cognitive and environmental sources of human error. Environmental sources include factors such as flawed interface design, confusing automation, and unexpected weather conditions. Cognitive sources include factors such as poor situation awareness, procedural non-compliance, and inadequate crew coordination. Many if not most significant incidents and accidents result from some combination of both cognitive and environmental factors. In fact, in a highly proceduralized domain such as aviation, with operators who are highly trained and well motivated, accidents rarely result from either environmental or cognitive causes alone. Training and experience are often sufficient to overcome even the most confusing interface designs, and the environment is often sufficiently redundant, reversible, and forgiving so that the vast majority of cognitive slips and mistakes have no serious consequences. Most highly consequential incidents and accidents result only when both environmental and cognitive factors collectively conspire to produce disaster.

Modeling to predict human error and its consequences therefore requires giving consideration to both cognitive and environmental issues. This report describes the results of a research project in which cognitive-environmental modeling, or more specifically pilot-vehicle-taxiway modeling, was performed in order to shed light on the possible causes of human error in aviation surface operations. Modeling consisted of integrating a pilot model developed within the ACT-R/PM cognitive architecture (Byrne & Anderson, 1998), a simplified model of aircraft dynamics during taxiing, and a visual model of the surface terrain, taxi routes, and signage. The work was performed under the auspices of the NASA Aviation Safety Program, Human Error Modeling element. The overall objective of that program element is to develop computational models with predictive capabilities to aid designers and analysts in identifying likely vulnerabilities in human-machine performance in aviation operations.

The particular modeling performed in this research was motivated by a set of experiments performed in NASA Ames' high fidelity Advanced Concept Flight Simulator (AFCS). Called T-NASA2 throughout this report, the scenario required 18 flight crews, consisting of active pilots from 6 commercial airlines, to approach, land, and taxi to gate at Chicago O'Hare International Airport (ORD) (Hooey, Foyle, Andre and Park, 2000). The primary purpose behind those experiments was to collect data for the evaluation of a suite of situation awareness and navigation aids currently being developed to both improve taxiing performance and to reduce the frequency and consequences of pilot error. The experiments contained both baseline conditions (current technology only) and conditions in which pilots were provided with various new display and communication technologies. The modeling performed in this research was focused solely on performance in the baseline (current technology) conditions.

2.2 Problem Definition: The T-NASA2 Data Set

Nine different taxiway routes were used in the baseline trials of the T-NASA2 simulation. Each of the 18 crews were tested over a balanced subset of 3 different routes for a total of 54 trials. Each trial began approximately 12 nautical miles out on a level approach into ORD. Pilots performed an autoland, and were required to taxi to the gate in simulated visually impoverished conditions (RVR 1000'). Further details can be found in Hooey, Foyle, Andre, and Parke (2000). It should be noted that the simulation did not represent all standard operating procedures (after landing checklists, log and company paperwork), nor all communication activities (with the cabin crew, dispatch, gate, etc.). As a result the level of crew workload was considerably less than a crew might experience in operational contexts (Goodman, 2001).

Across the 54 baseline T-NASA2 trials, a total of 12 off-route navigation errors were committed. On each, crews proceeded down an incorrect route without any evidence of immediate awareness, or else required correction by ground control. The NASA T-NASA2 research team designated these 12 to be "major errors." Additionally, 12 other deviations were observed but were detected and corrected by the crews themselves. These latter 12 deviations were thus classified as "minor errors" by the NASA research team. The NASA team provided our modeling team with detailed descriptions of each error, in terms of intersection complexity, turn type required, and a cognitive classification of each in terms of planning, decision making, or execution.

Two aspects of the T-NASA2 data set provided the primary motivation for the present modeling effort. First, it was believed that modeling might shed light on the underlying causes of the errors observed in the experimental simulations. A second motivation of the present modeling was the fact that the suite of situation awareness and navigation aids used in the new technology conditions of the T-NASA2 experiments were observed to eliminate navigation errors entirely. Our modeling approach therefore had as its goal to provide a cognitive explanation for the errors that were observed, in a fashion that was consistent with the finding that no errors were observed when the quality of perceptual information available for navigating was improved.

3. Modeling Approach

Modeling was based on three complementary analyses of the T-NASA simulation and the wider environment for airport taxi operations. The first, cognitive analysis, focused on creating an ACT-R/PM representation of the pilot navigating on the airport surface. The second, rational/ecological analysis, focused on describing the statistical regularities in the surface taxi environment generally, or at least with respect to a sample of 12 major U.S. airports. The third, environmental analysis, focused on creating a representation of the external entities with which the pilot model interacted in order to provide an accurate closed-loop model of human-machine system performance. These external entities included the aircraft itself, the visual information available to the pilot model, and the physical runway environment on which the simulated aircraft traveled. Each analysis resulted in a component of the overall computational model, which was integrated in such a fashion as to allow model behavior to be compared to human behavior observed in the T-NASA2 experimentation. Each of the three analyses and respective modeling components is described below.

3.1 Cognitive Analysis: The ACT-R/PM Cognitive Architecture

ACT-R/PM (Byrne & Anderson, 1998) augments the well-known ACT-R production system architecture with a perceptual-motor system realized as four perceptual-motor modules. Our preference for the ACT-R/PM version of the model reflects the simple fact that pilot behavior in aviation depends on perception/action interaction with the environment as well as on internal cognition. Central cognition is more or less serial (though spreading activation processes work in parallel) and each module is itself more or less serial, but the various components all run in parallel with one another. Thus, the production system could be retrieving something from long-term declarative memory while the Vision Module is shifting attention in the visual array and the Motor Module is preparing to press a key. This is in agreement with the original Model Human Processor of Card, Moran, and Newell (1983), which consisted of a collection of serial processors acting in parallel.

3.1.1 ACT-R Production System

ACT-R is a "unified theory of cognition" (Newell, 1990) designed to enable the modeling of a broad range of human behavior. The first four chapters of Anderson & Lebiere (1998) thoroughly describe ACT-R. ACT-R has three memories, a declarative memory containing chunks, which are facts like "3 + 4 = 7," a production (or procedural) memory containing production rules, IF-THEN condition-action mappings, and a goal stack also containing chunks, these encoding intentions. These memories are organized around the current goal, which is also a chunk.

ACT-R's behavior is centered around the production cycle. On each cycle, the activations of all memory elements are updated and production rules' IF sides are matched against the current goal and declarative memory. One of the productions matching its conditions is selected to fire. In ACT-R, only one production fires per cycle; conflict resolution is the process by which a single production is chosen when more than one production matches. Conflict resolution is based on a rational analysis (e.g., Anderson, 1990) of the expected utility of a production vs. its costs. Each production has associated with it a utility, defined as PG-C, where P represents the probability that the production will ultimately lead to accomplishing the goal, G represents the value of the goal, and C represents the costs incurred by the system of firing that particular production. Costs have generally been expressed in terms of time. In principle, given exposure to a particular environment, P and C will be learned by the system. Also, the utility computation is noisy so the system's behavior is not entirely deterministic. ACT-R's conflict resolution mechanism has been particularly important in the successful modeling of things like strategy choice (e.g. Lovett, 1998).

Productions generally create actions, such as changing the state of the goal or initiating physical action. Productions may also request the retrieval of a chunk from declarative memory, which plays an important role in accessing state information and problem-solving. The time to complete this operation depends on the activation of the chunk being retrieved with more active chunks being retrieved faster. Chunk activation, in turn, is a function of the estimated need odds of the chunk as computed by the following equation:

$$A_i = B_i + W_j S_{ji}$$

A_i represents the current activation of chunk i, B_i represents the base-level activation of chunk i, W_j represents the source activation spreading from chunk j, which must be a chunk referenced in the current goal, and S_{ji} is the strength of association between chunks j and i. The base-level activation of a chunk is a function of the usage history of the chunk and, if the chunk is not referenced, decays over time. Thus, chunk activations are dynamic and are a function of two things: relatedness to the current goal, and the history of use for that chunk in terms of frequency and recency. A substantial body of work has gone into the equations determining chunk activation values and dynamics (e.g. Anderson & Reder, 1999; Anderson, Bothell, Lebiere & Matessa, 1998). Chunk activations themselves are also noisy and thus provide another source of stochasticity in the system.

A subtlety that may not be immediately obvious from this description is that ACT-R is susceptible to a variety of workload effects. First, production firing in ACT-R is serial and this limits the effective bandwidth of the system. Second, the more items that ACT-R references in the current goal, the more diffuse the spreading activation will be and thus the more difficult it will be for the system to retrieve declarative knowledge. This diffusion of spreading activation has been used successfully to model multiple-task interference for tasks requiring memory retrieval (Byrne & Anderson, 2001). Furthermore, ACT-R can perform what is termed "partial matching." In essence, chunks that are similar to the chunk specified in the production for retrieval may be retrieved in place of the requested chunk. This is a function of the similarity of the two chunks in question and their activations. Chunks that are both very active and highly similar tend to generate partial matching, which is the source of errors in various ACT-R models of human memory performance (Anderson, Bothell, Lebiere & Matessa, 1998).

3.1.2 ACT-R/PM Modules

ACT-R/PM enhances ACT-R with perceptual-motor functionality to support modeling interactive behavior. Thus, ACT-R/PM allowed us to create simulation models of the entire human-environment system, representing cognition and behavior as a closed-loop process. The perceptual-motor enhancements in ACT-R/PM consist of four modules. The system is depicted in Figure 1.

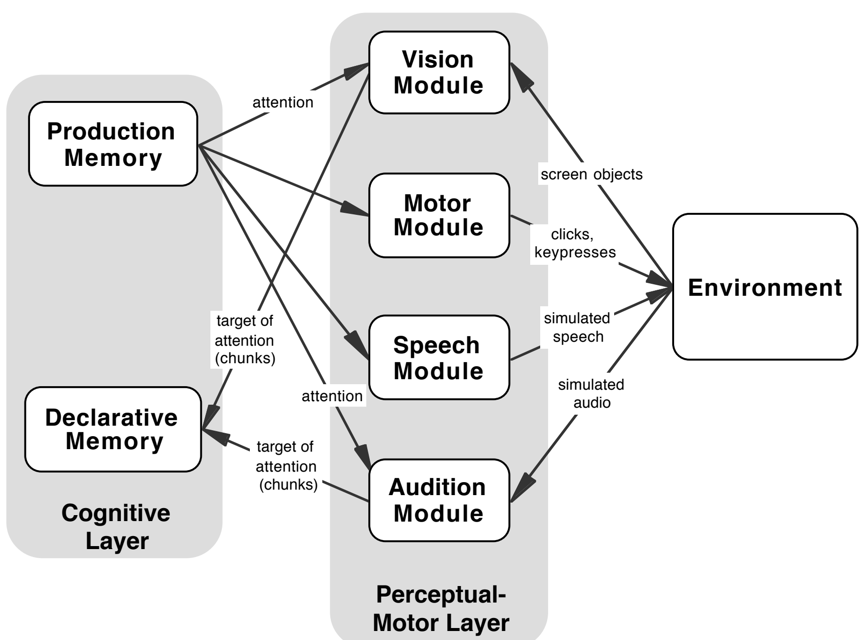

Figure 1. ACT-R/PM System Architecture

The *Vision Module* in ACT-R/PM is important in describing interactive tasks as it determines what the model "sees." A representation of the visual features present in the visual scene must be given to the Vision Module, and this is parsed to create an "icon." The Vision Module creates chunks based on the contents of the icon which provide declarative memory representations of the visual scene, which can then be matched by productions. For the Vision Module to create a chunk representing an object, visual attention must first be directed to the location of that object. Attention shifts happen asynchronously with respect to the production system and have a duration that is a system parameter that defaults to 135 ms, based on previous research (Anderson, Matessa, & Lebiere, 1997). This is again a limitation on bandwidth; only one object can be attended at once and the system must take time to shift.

ACT-R/PM's *Motor Module* is based directly on the description of EPIC's Manual Motor processor found in Kieras and Meyer (1996). The Motor Module represents ACT-R/PM's hands, and contains a number of parameters for representing movement. The Motor Module receives commands from the ACT-R productions to perform actions. In general, movement specification requires specification of a movement type, called a style, and one or more parameters, such as the hand/finger for making the movement. When a command is received by the Motor Module, the Motor Module goes through three phases: preparation, initiation, and execution. Timing for these phases is determined by empirical data (e.g. Rosenbaum, 1991) and movement models such as Fitts' Law.

The *Audition* and *Speech Modules* are at present somewhat less developed than their Vision and Motor counterparts. The Speech Module is somewhat rudimentary but it does produce timing estimates for the generation of simple speech. The Audition Module is somewhat parallel to the Vision Module in that it, too, is an attention-based system which can only be actively attending one signal at a time, again providing a bandwidth limitation; ACT-R/PM cannot effectively listen to multiple speech streams.

3.2 Rational/Ecological Analysis: The Taxi Navigation Decision

While the traditional domain of modeling in the history of ACT has been experiments in the psychology laboratory, it is clear that ACT-R's design is motivated by the desire to represent cognition and behavior as adaptive to more complex environments. But adaptivity is not defined in a vacuum; there is always the key question of what it is that the system adapts to and what strategies are available for adaptation at the knowledge level.

The approach we used for addressing this question is rational/ecological analysis, an examination of the environmental structure to which an agent must adapt. Rational/ecological analysis can conceivably inform the development of ACT-based models in several ways:

• Identification of procedures and problem-solving strategies used in a particular environment. This is essential in structuring the symbolic knowledge given to the model.

• Identification of the cost-benefit structure of those procedures and strategies, which is vital in setting conflict resolution parameters.

• Identification of the frequency, recency, associativity, and similarity structure of particular pieces of knowledge involved in the task. This is critical in the construction of the declarative knowledge base given to the model.

Rational/ecological analysis was our approach for providing ACT-R/PM with knowledge of the environmental structure to which T-NASA2 crews may have been adapted. Previous research led us to the view that some form of rational/ecological analysis would be necessary to faithfully describe pilot error in T-NASA2 in terms of limited adaptation. Casner, for example, described how pilot responses to clearances, and the control resources used to enact them, appeared to be adapted to the frequency with which specific clearances had been received during previous flights on the same routes (Casner, 1994). Pilots appeared to have strong expectations regarding the clearances they would receive and policies for using control automation that were attuned to the time available to respond to these clearances. Degani and his colleagues studied pilots' preferences for control modes as a function of a wide variety of environmental variables, and also how mode error could be seen as a reflection of deficiencies in environmental design (Degani, Shafto and Kirlik, 1999). Kirlik (1995) demonstrated how pilots' policies for using an autopilot seemed to be adapted to the costs incurred in programming the automation and the benefits achieved from doing so.

As modeling progressed, it became increasingly clear that a detailed analysis of the turn decision would play a key role in model development, and we thus focused our rational/ecological analysis efforts on this aspect of performance. Preliminary examination of taxiway visibility in relation to the taxi speeds observed in T-NASA2 and the simulated aircraft dynamics led us to the view early on that crews were likely to be supplementing the perceptual information obtained by looking out of the cockpit with some type of internally represented knowledge that allowed them to make correct turn decisions in the vast majority of situations even when visibility was highly limited. Goodman (2001) provided a detailed analysis of the geometrical properties associated with turn errors which seemed to indicate some trend towards increased error likelihood in cases where turns away from the terminal were required. All these factors led us to focus our rational/ecological analysis efforts toward an investigation of the task environment for the specific decision associated with which way to turn at a taxiway intersection.

A rational/ecological analysis was therefore performed to identify any general (i.e., airport neutral) regularities in the structure of the taxiway environment that might be internalized by flight crews through experience in these environments. Any such regularities could benefit pilots by giving them general strategies or expectations that could help them compensate for inadequate perceptual conditions such as those occurring in the T-NASA2 scenario. For this purpose, a study was performed with the aid of a Subject Matter Expert, a working airline (767) pilot, to assess any high level (i.e., airport neutral) knowledge pilots might use in deciding what turns to make during taxi operations. Jeppesen charts for all major U.S. airports were made available to the SME, who was asked to select charts for those airports for which he had significant experience of typical taxi routes. He selected 9 airports (DFW, LAX, SFO, ATL, JFK, DEN, SEA, MIA, ORD). The SME was asked to draw, using a highlighter on the charts themselves, the likely or expected taxi routes at each airport from touchdown to his company’s gate area. A total of 284 likely routes were generated in this way. This data supported our rational/ecological analysis of the turn decision environment. Results are discussed in section 4 of the report. These results played a significant role in guiding in the design of the turn decision ruleset embedded within our ACT-R/PM model, in particular, the Hierarchical Heuristic (HH) strategy in which turn decisions were based on cost/benefit considerations.

3.3 Environmental Analysis: Modeling the Visual and Controlled Elements

Environmental analysis concerns describing all the components of the human performance model representing entities outside the head of the human pilot. In this research, three such entities were modeled: the simulated aircraft controlled by the pilot model, the simulated visual information available to the pilot model, and the simulated runway and taxiway environment throughout which the simulated aircraft traveled. Each of these three environmental entities was computationally modeled and integrated with the cognitive components of the ACT-R/PM pilot model to create an overall representation of the interactive human-machine-environment system.

3.3.1 The Aircraft Model

Code for the vehicle dynamics that was used to drive the actual T-NASA2 flight simulator in which behavioral data was collected was unfortunately unavailable. We therefore had to create a simplified vehicle model with which the pilot model could interact. Given vehicle size, mass, and dynamics, however, we still did require a somewhat reasonable approximation to the actual aircraft dynamics used in the T-NASA2 experiments in order to be able to get a handle on timing issues. Although we were not interested in control issues per se, the dynamics of the aircraft played an important role in determining decision time horizons, a key factor in the cognitive representation of the pilot's activities.

Fortunately, we had access to resources to aid us in this aspect of the modeling. First, we contacted Dr. Dario Salvucci, then of Nissan Cambridge Basic Research, who generously supplied us with the code for his ACT-R/PM simulation of driver-automobile interaction (Salvucci, 2001). This code provided us with a template for how to link together the cognitive components of the human model and the engineering components of the vehicle model. In addition, it provided us with a baseline vehicle model we could then adjust to better simulate the taxiing dynamics of the aircraft used in T-NASA2 experimentation. The automobile model assumed that the driver controlled the vehicle in three ways: applying engine power, braking, and steering. For the purposes of modeling an aircraft during taxing, these three forms of control are sufficient, and needed only to be adjusted to mimic the T-NASA2 aircraft. Figures 2 and 3 below, from Cheng, Sharma, and Foyle's (2001) analysis of the NASA TSRV aircraft, show that we could proceed with a model in which it was reasonable to assume that throttle and braking inputs generated applied forces in a simple, linear manner.

We quickly learned, however, that the equations of motion for the Cambridge research automobile relied on a small angle approximation (for linearization) that was consistent with their own experimental scenarios but was inconsistent with the types of turns made in the T-NASA2 scenario. So, in addition to adjusting the acceleration and braking dynamics to mimic the simulated aircraft, we had to either write our own equations of motions for a taxiing aircraft during turning, or else develop another solution to this problem. We chose the latter.

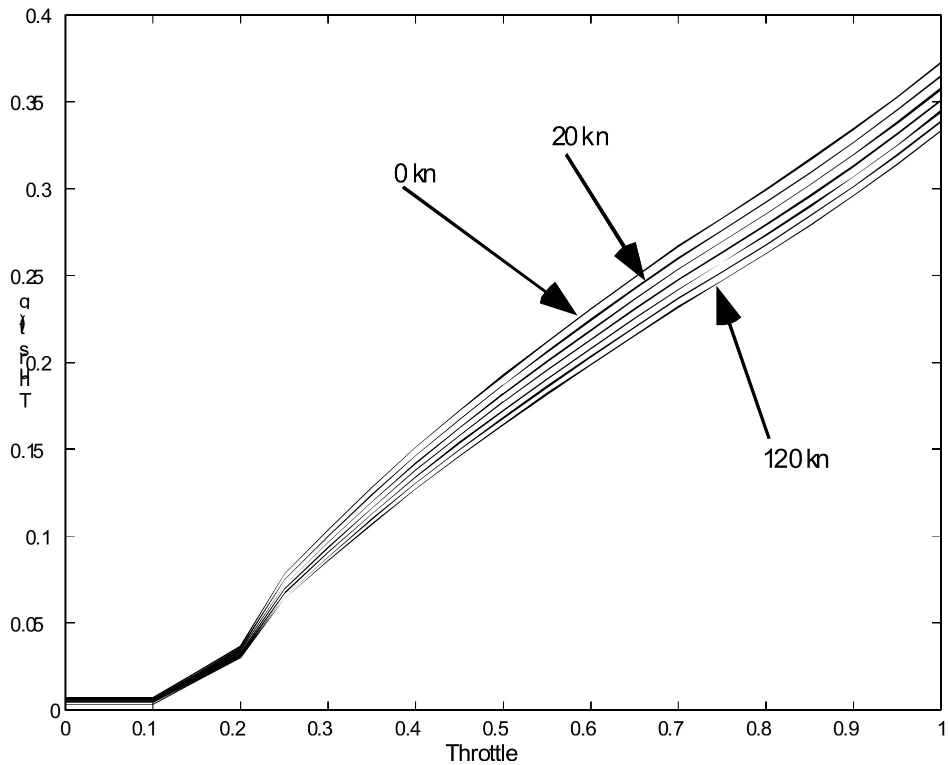

Figure 2. Linear effects of throttle on thrust at various speeds (from Cheng et al. 2001)

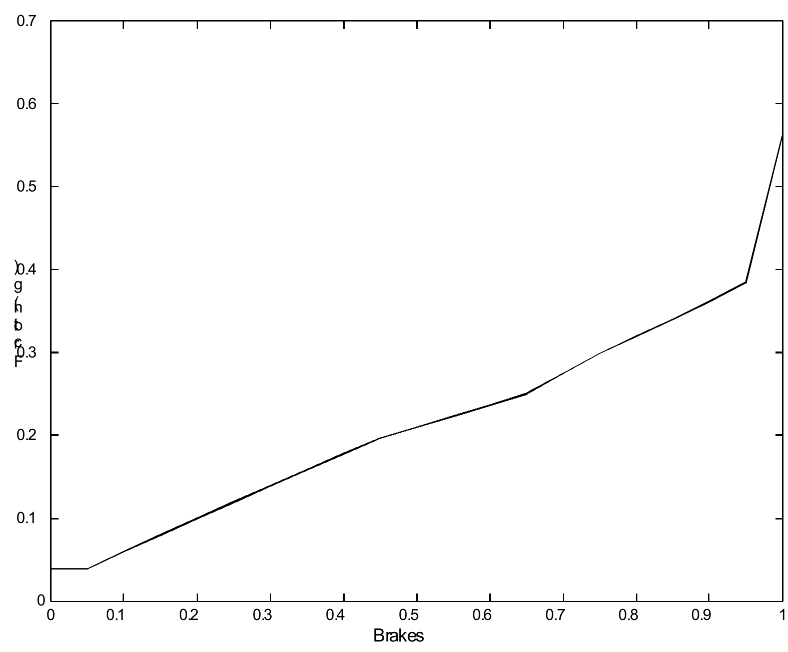

Figure 3. Linear effects of braking on friction force in g's (Cheng et al. 2001)

After consideration of the functional role that steering inputs play in the T-NASA2 scenario, we decided that we could finesse the problem of steering dynamics by assuming that the manual control aspects of the steering problem did not play a significant role in the navigation errors that were observed. That is, we assumed that making appropriate turns was purely a decision making problem, and that no turn errors resulted from correct turn decisions that were erroneously executed. Note that this assumption does not completely decouple the manual and cognitive aspects of the modeling, however. It was still the case that manual control of the acceleration and braking aspects of the model did play a role in determining the aircraft's position relative to an impending turn, and its speed of approach to the turn. These variables affect the time available to make a turn decision, and in our implementation, as this time is reduced there is a greater probability of an incorrect turn decision. Our simplification regarding steering merely boils down to the fact that once the model has made its decision about which turn to take, that turn is then executed without error. We believe that it may be possible to model some errors at the level of turn execution in ACT-R/PM, in particular, mapping from a desired turn direction to the selection of a particular yellow line on the taxiway, and then the mapping of that yellow line to specific manual control commands, are not trivial operations for ACT-R/PM. We suspect that some of the execution errors are a result of difficulties with one or both of these mappings. However, given our time constraints this did not seem to be the most fruitful approach relative to modeling decision errors.

To implement this aspect of the model, we decided to model the ORD airport taxiway as a set of interconnected "rails" upon which travel of the simulated aircraft was constrained. Taxiway decision making in this scheme, then, boiled down to the selection of the appropriate rail to take at each taxiway intersection. In this manner, we did not have to model the dynamics of the aircraft while turning: we simply moved the aircraft along each turn rail at a specified, nominal, speed. This speed was, however, influenced by the severity of the turn, in order to provide a reasonable approximation to the T-NASA2 data. To determine this speed, a simple analysis was made to determine the maximum speed at which the aircraft could proceed in a turn given the turn's radius and the constraint that lateral acceleration was limited to .25 g (Cheng et al. 2001; see also the turn speed data reported in Cassell, Smith, and Hicok, 1999). The results of that analysis are depicted in Figure 4.

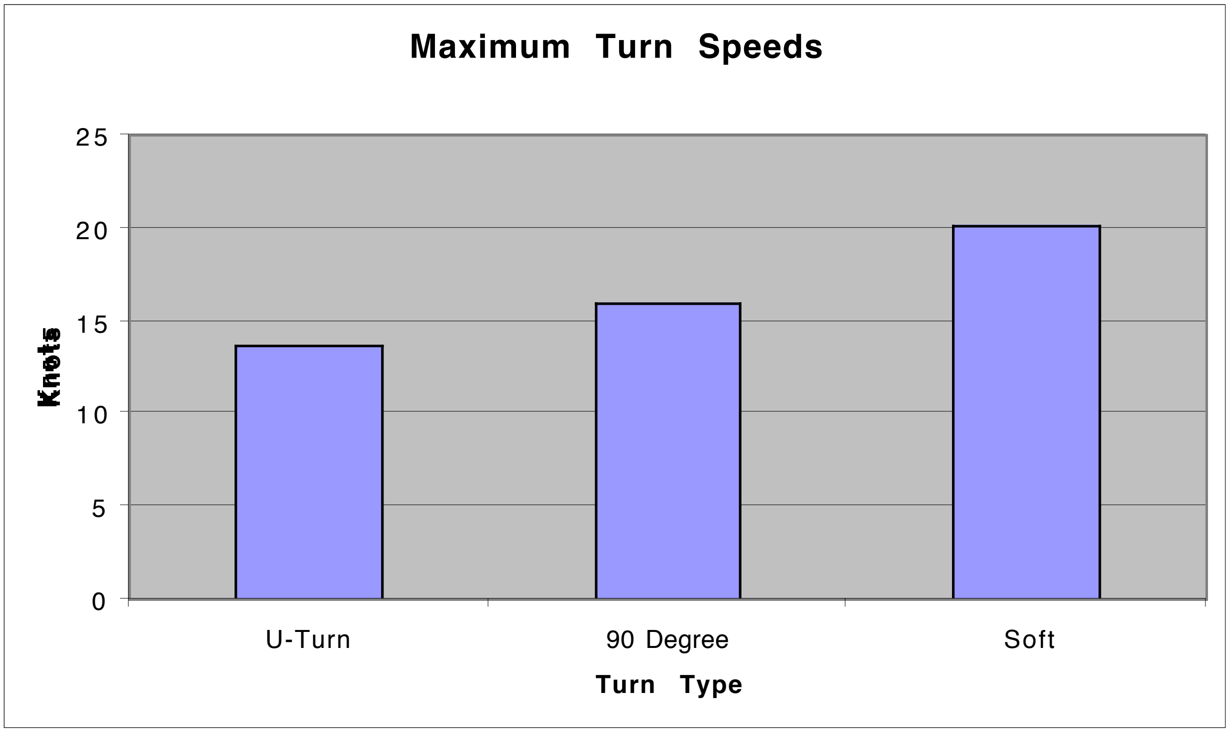

Figure 4. Maximum turn speeds used in the model as a function of turn type.

3.3.2 The Visual Scene Model

The model used to represent the visual information available to our ACT-R/PM pilot model was obtained from the actual T-NASA2 flight simulator in the form of a software database. This database consisted of location-coded objects (e.g., taxiways, signage) present on the ORD surface, or at least those objects presented to flight crews during T-NASA2 experimentation. Distant objects became "visible" to the pilot model at similar distances to which these same objects became visible to human pilots in T-NASA2 experimentation. Construction of this model from the raw polygon data provided by NASA was not trivial.

3.3.3 The Runway/Taxiway Model

As discussed in section 3.1.1, we modeled the runway and taxiway environment which actually constrained the motion of the simulated aircraft in terms of a network of interconnected "rails." This model describing the environment on which the piloted aircraft can actually travel should be kept distinct from the visual representation of that environment as discussed above in section 3.3.2. The actual locations within the simulated environment that the aircraft could potentially occupy were but a small subset of the locations depicted in the visual scene model. The visual scene model contained information specifying the layout of the airport surface, whereas the runway and taxiway model contained information that actually constrained (and in turns, partly generated) the motion of the simulated aircraft. This runway and taxiway model was computationally implemented as linked lists of rails designating paths of potential travel.

4. Analysis and Modeling Results

Previous sections have described the approach that was used to analyze and model the relevant entities in the dynamic, interactive, pilot-aircraft-visual scene-runway/taxiway system. In this section, we present the results of the ecological/rational analysis of the turn decision, and how the results of this analysis guided the development of the ACT-R/PM pilot model. The overall structure of the pilot model, as well as how this model interacted with the aircraft, visual, and runway/taxiway models is described.

4.1 Results of the Rational/Ecological Analysis of the Turn Decision

As described in section 3.2, we conducted a study of the turn decision task environment using an SME as an experimental participant to provide data for a rational/ecological analysis of this decision. Jeppesen charts for all major U.S. airports were made available to the SME, who was asked to select charts for those airports for which he had significant experience of typical taxi routes. He selected 9 airports (DFW, LAX, SFO, ATL, JFK, DEN, SEA, MIA, ORD). The SME was asked to draw, using a highlighter on the charts themselves, the likely or expected taxi routes at each airport from touchdown to his company's gate area. A total of 284 likely routes were generated in this way.

The data gathered in this study informed our analysis of the array of potential decision strategies available for making turn decisions. We identified four possible strategies, presented below in increasing order of accuracy:

1. Remember the correct route: While fast, this strategy is increasingly inaccurate as time lapses between obtaining the list of turns described in the clearance and the time at which turn execution is actually required.

2. Make turns toward the gate: While somewhat slower than the first strategy, this strategy has a reasonable level of accuracy at many airports.

3. Turn in the direction that reduces the larger of the X or Y distance between the aircraft and the gate. We deemed this strategy to be moderately fast, like strategy 2, but with an even higher level of accuracy than strategy 2.

4. Derive from map/spatial knowledge. This is the slowest strategy available, with high accuracy possible only from a highly experienced (at a given airport) flight crew.

5. Buy time and reassess by braking: This strategy is not a turn strategy per se, but rather a technique for creating more time for the execution of one of the above strategies.

We were somewhat surprised how effective the "quick and dirty" heuristic strategies described in 2 and 3 above turned out to be at the airports studied. Figure 5 presents the results of an analysis of the effectiveness of these two heuristic strategies. Note that the XY heuristic (3) is

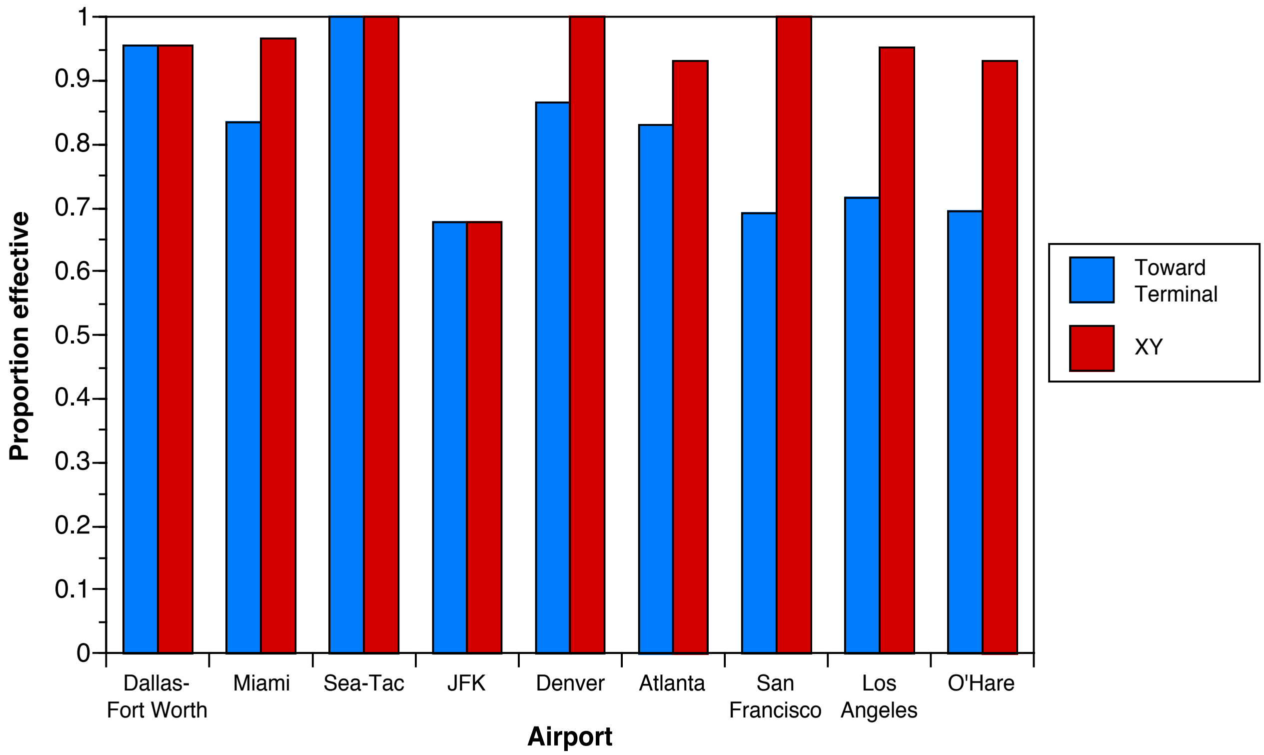

Figure 5. Heuristic Effectiveness by Airport Studied

quite good across the board, and the even simpler "toward terminal" heuristic is good a many airports (but not O'Hare). It is noteworthy that at all turns simulated in the T-NASA2 scenario at which both these heuristics fail, flight crews were observed to make at least one error. We take this finding to warrant our assumption that, due to impoverished perceptual conditions in the T-NASA2 scenario, pilots were supplementing whatever perceptual information they could obtain with general, airport-neutral knowledge of the likely clearance patterns to be given to travel a path from touchdown to gate. As such, we created the turn decision making components of the ACT-R/PM pilot model to make decisions according to the set of 5 heuristics described above. One can think of these heuristics as being hierarchically organized in terms of their costs and benefits. The pilot model works by attempting to use the strategy that achieves the highest effectiveness given the time available. In some cases it will apply the brakes to increase the time available for decision making as well, shaping, in a fashion, its own decision environment.

Figure 6 depicts average taxi route length, measured as the number of turns, at each of the airports studied. We speculate that the significantly longer routes required at O'Hare might also contribute to its reputation as a relatively difficult-to-navigate airport surface.

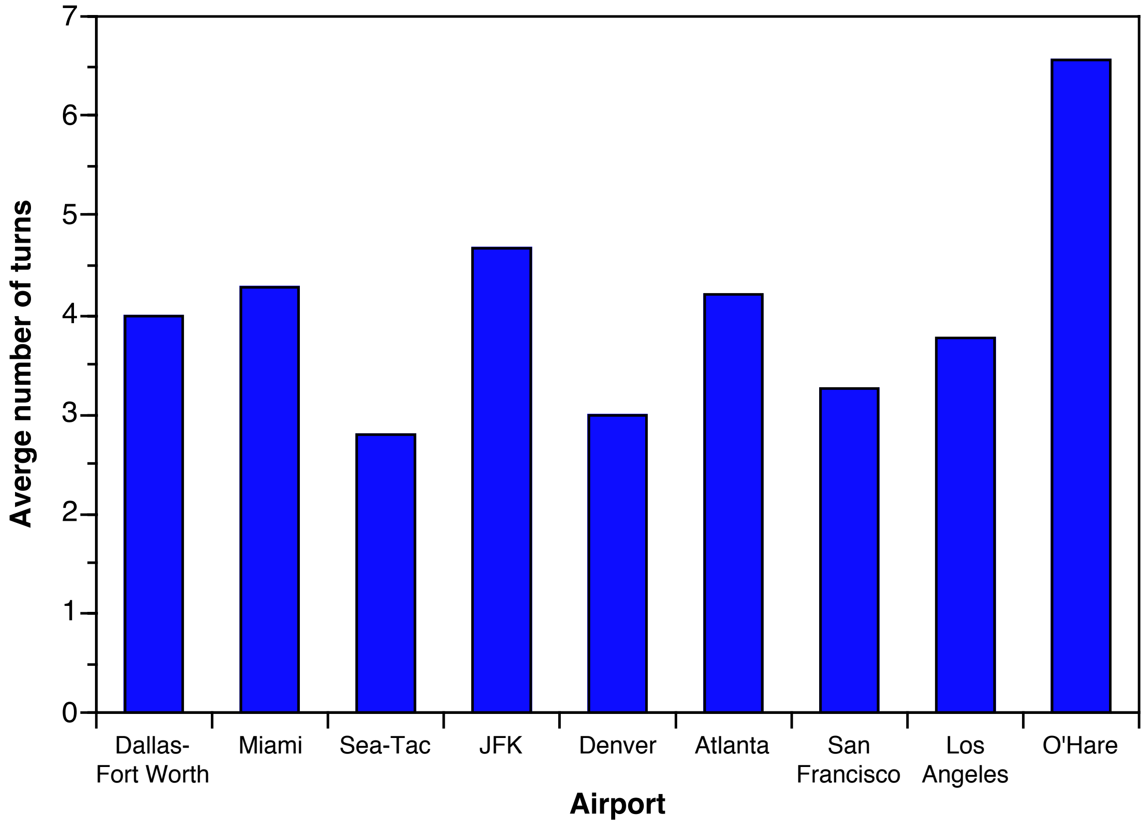

Figure 6. Average Taxi Route Length by Airport Studied

4.2 The ACT-R/PM Pilot Model

The ACT-R/PM pilot model was designed with productions implementing the results of the rational/ecological analysis of the turn decision, as described above. In addition, the model was computationally coupled to the aircraft, visibility, and runway/taxi components of the overall model that represented the environment external to the pilot model. This allowed the pilot model to run in a closed-loop fashion, simulating the dynamic, interactive relationship between the pilot, his vehicle, and his surroundings during the T-NASA2 scenario.

4.2.1 Model Scope

A number of simplifications had to be made to achieve an implementation of the model in the time allowed. The current implementation models the pilot but not the first officer. Also, there is no attempt to model any errors resulting from miscommunications between agents (e.g., among the flight crew, between the flight crew and air traffic control). Finally, recall that, due to the limitations in the vehicle model, the manual control activities associated with steering are not modeled. The model does, however, respect the g-force constraints associated with steering around curves of various radii, and does devote attentional resources to simulate the perceptual and motor demands associate with actual steering.

4.2.2 Overall Decision Flow

Figure 7 depicts the overall decision, or goal selection, flow of control in the ACT-R/PM pilot model. Just as with decision flow diagrams of human subjects, we do not intend to imply that this graphical flow diagram is actually tokened in the computational model as such. A set of rule-based productions, managed through a goal stack as described in section 3.1, is actually responsible for implementing the flow of control illustrated in Figure 7. The graphical depiction is provided here solely for clarity of description.

The top goal, taxi to the gate, is by design the overarching goal to which control returns when lower level goals have completed processing. During routine (straight) taxiing, the model cycles through the four "maintenance" goals. When one of these goals completes, it can return information to the top goal. For example, if an incursion were to be detected, the completed subgoal will return a note to the main goal, which will cause the main goal to push a subgoal to actually handle the incursion. In another case updating the aircraft location can result in identifying that an intersection is coming up soon, which will return a note to the main goal to deal with the decision making associated with navigating the intersection. One important aspect of this cycle is that each of the maintenance goals takes time, and thus the model may not immediately be in position to deal with an upcoming intersection the instant information about that intersection becomes available.

4.2.3. Look for Incursion

This purpose of this goal is to look for anything unusual on the rail currently be traveled. This goal will also pick up other relevant information, such as signs that are in view. If there is no incursion detected, control is returned to the top goal, and if there is an incursion, the top goal is informed of such.

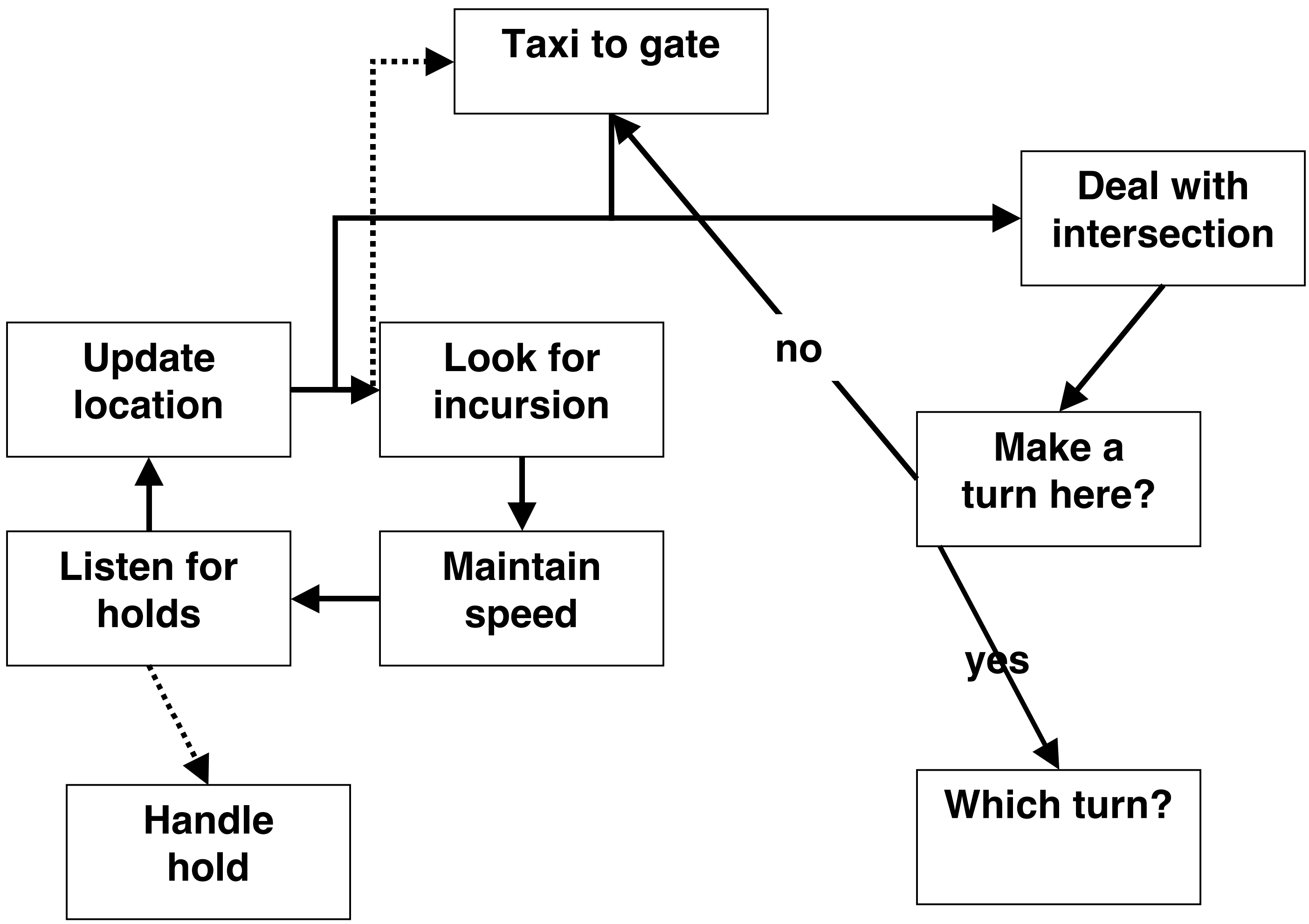

Figure 7. Decision (Goal Selection) Flow in the ACT-R/PM Pilot Model

4.2.4 Listen for Hold

Another possibility in this task environment is that ground control will instruct a flight to hold position along the route. Thus, the model must devote some time to active processing of the audio stream (if any) to determine if a hold instruction has been issued. This maintenance goal can, of course, return a note to the main goal that a hold instruction has been received, which causes the main goal to push a new subgoal to handle the hold. Note that we did not actually implement the "handle hold" procedure as only the active checking for hold instructions actually impacts the timing of turn decisions.

4.2.5 Maintain Speed

Since the model does not perform manual control of steering, this goal is relatively simple. Speed is checked against nominal lower and upper speed bounds. If the aircraft is moving too fast, then depending on the situation either the throttle is backed off or braking is applied. If the aircraft is moving too slowly, then either braking is backed off or thrust is applied. Again, time is required for each of these activities.

4.2.6 Update Location

The location of the aircraft is represented within the pilot model in a qualitative way (e.g., on taxiway X, between taxiways Y and Z, heading toward taxiway Z). The qualitative location is updated by reference to signs and the visible intersections of taxiways themselves.

4.2.7 Make a Turn Here?

A simple logic is used here. In some cases, the taxiway geometry makes determination of whether or not to make a turn trivial. For example, if the upcoming intersection is a T, then a turn must be made. Otherwise, the model generally relies upon memory of turns as represented in the clearance to know whether or not a turn should be made (but not necessarily in which direction). Our model implements the memory for this list much like the serial memory model presented in Anderson, Bothell, Lebiere, & Matessa (1998) and is thus subject to similar errors. However, in the T-NASA 2 error corpus, only one of the six decision errors was of this type, so we did not focus on this as a primary source of error behavior.

4.2.8 Which Turn?

This goal was the focus of the rational/ecological analysis of turn decisions, and is described in detail in section 4.1. The hierarchical heuristic (HH) strategy was used to select a turn decision making heuristic based on temporal and accuracy cost/benefit considerations.

4.3 Model Output

ACT-R/PM outputs a complete trace of behaviors at a very small grain size: a millisecond-by-millisecond account of the behavior being modeled. Of course, for scenarios at the time scale represented by taxiing, this resolution is somewhat of an overkill. At the present time the overall model, including pilot, aircraft, visual, and runway/taxiway components is implemented and can be run in a dynamic, interactive fashion.

In addition, we will not be running the model only once. ACT-R/PM's behavior is both dynamic and non-deterministic. Workload that is generated during a model run can impact local behavior at any particular time, and the downstream results of small fluctuations are not always predictable. Thus, we expect to conduct a number of Monte Carlo simulation runs and both analyze the results of particular runs to see if we can isolate factors that led to errors made by the model on any particular run as well as aggregating across the Monte Carlo runs to extract general trends. However, the process of performing Monte Carlo runs of the model continues.

5. Findings, Lessons Learned, and Recommendations

Findings, lessons learned, and recommendations resulting from this research fall into two general categories: technical issues in computational cognitive modeling, and the interpretation of T-NASA2 data.

5.1 Technical Issues in Computational Cognitive Modeling

As this report hopefully demonstrates, it is a nontrivial matter to apply scientific models of cognition, developed and validated primarily with laboratory data, to applied contexts such as human-machine performance in aviation. Specific challenges include the following.

5.1.1 Specification of Model Inputs and Outputs

Experimental tasks are typically carefully designed in such a way that the inputs to, and outputs of, the cognitive system are readily identifiable. This is largely done by making the perceptual and motor demands associated with cognitive experimentation relatively trivial. Unfortunately, the perceptual and motor demands associated with aviation cognition can be extensive. This problem surfaced in this research in the difficulties associated with coupling the ACT-R/PM model with the visual scene database, and also to some extent, with the aircraft model. The latter was not a severe problem because motor outputs could be considered to be relatively discrete in this instance, consisting of discrete settings of the throttle and brake. Modeling cognitive activity in the presence of a richer set of motor outputs would present additional challenges.

5.1.2 Modeling Environmental Objects and Dynamics

As this report hopefully demonstrates, achieving a reasonable model of pilot cognition in dynamic, interactive contexts can depend on the availability of reasonable models of the visual, physical, and controlled environment, as well as its dynamics. The dynamics of human cognition and behavior is interleaved with, and in concert with, the dynamics of environmental entities that also participate in the functioning of the integrated human-environment system. While modeling these external objects and dynamics is not properly a problem for psychology, but instead a matter for the applicable content area (e.g., aeronautical engineering), modeling these dynamics in a manner so that they can function interactively with a cognitive model is a matter for psychology. Cognitive modeling software packages can make better and more explicit provisions for representing objects and dynamics in the external environment so facilitate the task of modeling interactive behavior in contexts more complex than the desktop computer. Again, this is not a trivial problem; machine vision researchers have been working on problems in dynamic vision for decades and still lack clear solutions to many of the problems faced by cognitive models that interact with moving environments.

5.1.3 Timing Issues

Operating systems for running human-environment system models should provide separate clocks and processing resources for simulating cognitive and environmental dynamics. Neither is subservient to, nor should be subsumed, by the other, nor should either model have to wait while the other is updating. The current approximation in which processing resources are passed back and forth between cognitive and environmental components will be found to be increasingly unwieldy as more dynamic contexts are modeled.

5.2 Interpretation of T-NASA2 Data

Given the constraints of the modeling architecture described above, we are encouraged by the results of this research to continue to pursue ACT-R/PM models of human performance in dynamic, aviation contexts. We believe that the human errors observed in the T-NASA2 scenario were consistent with the results of our rational/ecological analysis of the turn decision. As such, we believe that the view of cognition embodied in ACT-R/PM as constrained adaptation to the statistical and cost-benefit structure of the previously experienced task environment achieves some level of support from this research.

The crux of the interpretation of human errors in T-NASA2 is that a performer has multiple, redundant methods for handling individual decisions, and will use the most accurate strategy possible given the time available. When time is short, or perceptual conditions do not supply unambiguous specific information, the decision maker will tend to rely on computationally cheaper, but less specific information gained from experience with the wider class of decision situations of which the current decision is a member. In the case of the T-NASA2 scenario, this more general information pertained to the typical taxi routes that would be expected from touchdown to gate at major U.S. airports. We do not believe that it was coincidence that errors in the T-NASA2 scenario were nearly always made at points where the correct clearance indicated a turn in a direction that would be considered unexpected or unusual given the typically expected routes of travel between touchdown and gate. This particular error mode is very much in the spirit of the ACT-R/PM modeling approach (which views all performance as constrained adaptation), and it could be represented in the model in a fairly straightforward, rule-based manner. This interpretation is also consistent with the fact that the suite of display aids used in the high-technology conditions of T-NASA2 experimentation, by providing improved perceptual information, effectively eliminated human errors. Given this modeling demonstration, we hope that this research will convince more members of the human factors and aviation safety community to investigate human performance issues with the benefit of emerging developments in computational modeling

6. Acknowledgments

The authors would like to thank Captain Bill Jones who lent his considerable expertise to this project by serving as Subject Matter Expert. We would also like to acknowledge the important modeling contributions of Brian Webster, Michael Fleetwood, and Chris Fick of the Psychology department at Rice University. Finally, we are grateful for the support of the NASA Ames T-NASA2 research team, who generously provided their time, data, and expertise to this project. We would especially like to thank Allen Goodman, Lynne Martin, James Johnston, Dave Foyle, and Tina Beard for their patient support on a week by week, and sometimes even day by day, basis.

7. References

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Lawrence Erlbaum.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Bothell, D., Lebiere, C. & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language, 38,* 341–380.

Anderson, J. R., & Lebiere, C. (1998) *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

Anderson, J. R., Matessa, M., & Lebiere, C. (1997). ACT-R: A theory of higher level cognition and its relation to visual attention. *Human-Computer Interaction, 12*, 439–462.

Anderson, J. R., & Reder, L. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General, 128,* 186-197.

Bisantz, A., Kirlik, A., Gay, P., Phipps, D., Walker, N., and Fisk, A.D. (in press). Modeling and analysis of a dynamic judgment task using a lens model approach. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans*.

Bisantz, A., and Kirlik, A. (2000). Adaptivity and rule verification: An empirical investigation. *International Journal of Cognitive Ergonomics*.

Byrne, M. D. (2001). ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies, 55,* 41-84.

Byrne, M. D., & Anderson, J. R. (1998). Perception and action. In J. R. Anderson & C. Lebiere (Eds.) *The atomic components of thought* (pp. 167–200). Mahwah, NJ: Lawrence Erlbaum.

Byrne, M. D., & Anderson, J. R. (2001). Serial modules in parallel: The psychological refractory period and perfect time-sharing. *Psychological Review, 108,* 847-869.

Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science, 21,* 31–61.

Card, S., Moran, T., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum.

Cassell, R., Smith, A., and Hicok, D. (1999). Development of airport surface required navigation performance (RNP). NASA/CR-1999-209109.

Casner, S. (1994). Understanding the determinants of problem solving behavior in a complex environment. *Human Factors 36(4),* 580-96.

Cheng, V. H. L., Sharma, V., and Foyle, D. C. (2001). Study of aircraft taxi performance for enhancing airport surface traffic control. Manuscript accepted for IEEE-ITS publication.

Degani, A., Shafto, M. & Kirlik, A. (1999). Modes in human-machine systems: Review, classification, and application. *International Journal of Aviation Psychology, 9(2),* 125–138.

Goodman (2001). Personal communication.

Hooey, B. L., Foyle, D. C., Andre, A. D., and Parke, B. (2000). Integrating datalink and cockpit display technologies into current and future taxi operations. *Proceedings of the AIAA/ IEEE/SAE 19th Digital Avionics System Conference*.

Kieras, D. E., & Meyer, D. E. (1996). The EPIC architecture: Principles of operation. Unpublished manuscript from ftp://ftp.eecs.umich.edu/ people/kieras/EPICarch.ps

Kirlik, A. (1993). Modeling strategic behavior in human-automation interaction: Why an "aid" can (and should) go unused. *Human Factors, 34(2),* 221-242.

Kirlik, A., Miller, R.A., and Jagacinski, R.J., (1993) Supervisory control in a dynamic uncertain environment: A process model of skilled human-environment interaction. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol 23, No. 4, (pp. 929-952).

Lee, F. J. (2000). Does learning of a complex task have to be complex? A study in learning decomposition. Ph.D. Thesis, Carnegie Mellon University. Pittsburgh, PA.

Lovett, M. (1998). Choice. In J. R. Anderson & C. Lebiere (Eds.) *The atomic components of thought* (pp 255–298). Mahwah, NJ: Erlbaum

Rosenbaum, D. A. (1991). *Human motor control*. New York: Academic Press.

Salvucci, D. (2001). Predicting the effects of in-car interfaces on driver behavior using a cognitive architecture. In *Human Factors in Computing Systems: CHI 2001 Conference Proceedings* (pp. 120-127). New York: ACM Press.

Schoelles, M. J., & Gray, W. D. (2000). Argus Prime: Modeling emergent microstrategies in a complex, simulated task environment. Paper presented at the 2000 International Conference on Cognitive Modeling.