Modeling Eye-Movements in a Dynamic Task

Frank J. Lee (fjl+@cmu.edu)
John R. Anderson (ja+@cmu.edu
Carnegie Mellon University
Department of Psychology
Pittsburgh, PA 15213
USA

Abstract

In this paper, we present an ACT-R/PM (Byrne & Anderson, 1998) model of skilled performance in the Kanfer-Ackerman Air Traffic Controller Task©. With the model, we attempt to account for the latencies and the eye-movement patterns of the participants performing this task.

Introduction

Many unified theories of cognition are beginning to incorporate a comprehensive theory of the perceptual-motor system within their architectural framework. For example, Chong and Laird (1997) have developed EPIC-Soar that combines Soar (Newell, 1990) with the perceptual-motor components of EPIC (Meyer & Kieras, 1997). And Byrne and Anderson (1998) have developed ACT-R/PM that combines ACT-R (Anderson & Lebiere, 1998) with the perceptual-motor components of EPIC and the theory of visual attention from the Visual Interface (Anderson, Matessa, & Lebiere, 1998).

Endowed with a detailed theory of the perceptual-motor system, these extended architectures like EPIC-Soar and ACT-R/PM can now make detailed predictions about the perceptual-motor processes, including eye-movements. In this paper we present an ACT-R/PM model of skilled performance in the Kanfer-Ackerman Air Traffic Controller (KA-ATC) Task® (Ackerman, 1988, Ackerman & Kanfer, 1994). We compare the performance and the eye-movement behavior of the model to those of the participants and show how eye-movement data can be invaluable (and often necessary) in developing and evaluating a model. Before we begin, we first provide a brief description of KA-ATC task.[1]

Kanfer-Ackerman Air Traffic Controller Task©

A typical display of the KA-ATC task is presented in Figure 1. As can be seen, the KA-ATC task is composed of the following display elements: (a) twelve hold pattern positions divided into three levels, (b) four runways, (c) feedback information for participants on their current

[1] For a more complete description of KA-ATC task, please consult Ackerman & Kanfer (1994).

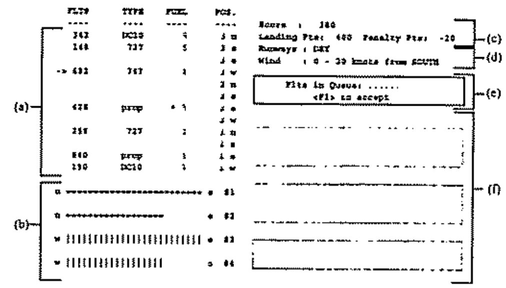

Figure 1. A prototypical display of the Kanfer Ackerman Air-Traffic Controller Task©.

performance, (d) current conditions of the runways, wind direction, and wind speed, (e) a queue of planes waiting to enter the hold pattern, and (f) three message windows.

Three principal actions that people can perform this task are (1) accepting planes from the queue into a hold pattern, (2) moving planes within the three hold levels, and (3) landing planes on a runway. These actions are accomplished by using a combination of four keys: ↑, ↓, F1, and ↵. The ↑ and the ↓ keys move the cursor up and down between the hold positions and runways, the F1 key accepts the planes from the queue into a holding position, and the ↵ key can select a plane in the hold, place a selected plane (either from the queue or from another hold position) into an empty hold position, or land a plane on the runway.

Score is calculated as follows: 50 points for landing a plane, 100 points deducted for crashing a plane, and 10 points deducted for violating a rule. Participants also lose 10 points for each minute that a plane's fuel is under 4 minutes at the time of the landing. When the fuel level of a plane falls to 0 minutes, it crashes. That is, 100 points are deducted from the score, and the crashed plane is removed from the hold position.

In addition, planes enter the queue approximately every 7 seconds and a plane takes 15 seconds to taxi across a runway. Since only two runways can be used at a time (depending on wind direction) people can never exhaust the planes accumulating in the queue over the course of the experiment. However, since planes in the queue also do not lose fuel, there is no real pressure to empty out the queue. But, once the planes are entered from the queue into the hold levels, they have between 4 to 6 minutes of fuel and begin to lose fuel in real time. Hence, they need to be attended to (i.e. landed) quickly to avoid a low fuel penalty or even crashing them when the fuel level goes to 0 minutes.

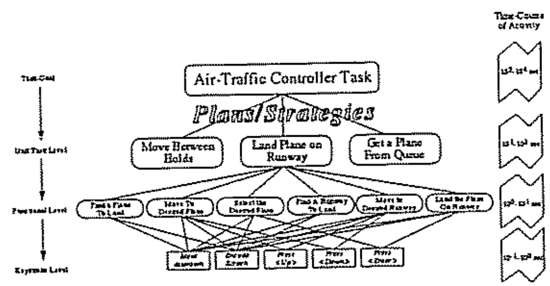

Figure 2. A decomposition of the KA-ATC task.

Task Analysis

Figure 2 illustrates our decomposition of the KA-ATC task. As can be seen, the overall task can be decomposed into performing a sequence of three unit tasks that correspond to the three principle actions used in this task. Each unit task can be further decomposed into a number of functional-level subgoals. For instance, the unit-task of landing a plane involves (1) finding a plane to land, (2) moving to the plane, (3) selecting the plane, (4) finding a runway to land, (5) moving to the desired runway, and (6) landing the plane. Each of these functional-level subgoals involves a number of keystroke-level subgoals, including a sequence of shifts of attention across the screen, encoding of information on the screen, and a keystroke to effect the desired action. In the next section, we use our task decomposition as the basis for our data analysis and our ACT-R/PM model of the task.

Data and Model

ACT-R/PM (Byrne & Anderson, 1998) is an extension to ACT-R (Anderson & Lebiere, 1998) that adds a detailed theory of perceptual and motor action to the ACT-R cognitive architecture. ACT-R/PM integrates ACT-R with the theory of visual attention from the Visual Interface (Anderson, Matessa, & Lebiere, 1998) and the theory of the perceptual-motor system underlying the EPIC (Meyer & Kieras, 1997) architecture. The strengths of ACT-R/PM lie with its ability to provide parallel execution across the cognitive system and the perceptual-motor system and with its ability to provide a detailed control over the perceptual-motor system.

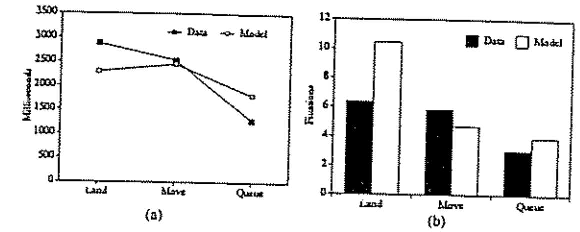

Figure 3. (a) The mean time to complete a unit tasks for the participants versus the model, and (b) the mean number of fixations used by the participants and the mean number of attention shifts made by the model during a unit task execution.

In this section, we present an ACT-R/PM model of the Trial 18 performance of the participants from Lee (1999). In Lee (1999), people performed 18 trials of the KA-ATC task with each trial lasting 10 minutes. In addition to various performance measures, i.e. score, their eye-movement protocols were collected.

Performance at Trial 18

Since there were 10 subjects in our eye-tracking experiment, we ran our expert model of the KA-ATC task for 10 iterations. At Trial 18, the participants on average landed 73.8 (S.D. = 5.49) planes, while the model on average landed 76.7 (S.D. = 2.35) planes. The greater variance in the number of planes landed by the participants compared to the model reflects the greater diversity in the strategies used by the participants (John & Lallement, 1997; Lee, Anderson, and Matessa, 1995; Reder & Schunn 1998; Schunn & Reder, submitted). The model, on the other hand, employs a single efficient strategy.

Unit Tasks

In Figure 3, we plot (a) the mean time to complete a unit tasks for the participants and the model, and (b) the mean number of fixations used by the participants and the mean number of attention shifts made by the model during a unit task execution. To measure the fixations of the participants, we used a velocity-based algorithm (Douglass, 1998) to isolate fixations from saccades. And to measure the attention shifts made by the model, we counted the number of times that the model fires a production that shifts its attention across the task screen.

As can be seen in Figure 3 (a), the participants on average require more time to land a plane than the model, while they require less time to queue a plane then the model. However, both the participants and the model require about the same amount of time to move a plane.

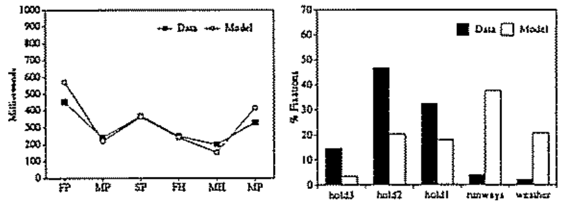

Figure 4. (a) The mean time to complete the functional-level subgoals of the move unit task for the participants and the model, and (b) the distribution of the participants' fixations and the model's attention shifts across the task display during a move unit task.

The differences in the time to complete a land and the queue unit task between the participants and the model reflect the greater number of strategies used by the participants in selecting a unit task. The number of keystrokes for the same type of a unit task can vary widely and are particularly sensitive to the queuing strategy used by the participants (e.g. John & Lallement, 1997). Since the participants vary widely in their queuing strategy whereas the model follows a single efficient queuing strategy, one can almost expect that there will be a difference in the unit task completion times. However, this is not the case with the functional-level keystrokes, which we will examine next. Functional-level keystrokes reflect the time to complete a single keystroke and hence are less sensitive to the differences in strategy use.

As can be seen in Figure 3 (b), the number of fixations that the participants make is similar to the number of attention moves made by the model in both the move unit task and the queue unit task. But for the land unit task, the model makes four additional attention moves compared to the participants. The model makes these additional attention moves, because ACT-R/PM currently lacks the ability to process peripheral information. The model, therefore, needs to repeatedly attend to the runways to figure out when they have cleared. The participants, on the other hand, are able to acquire this information peripherally, and only need to attend to the runways when they have cleared. Hence, the poor fit between the model and the participants for the land unit task in Figure 3 (b) reflects ACT-R/PM's the current inability to respond to peripheral information.

Move Unit Task

Of the three unit tasks, we focus on the move unit task for a detailed analysis. Figure 4 (a) plots latencies of the functional-level subgoals of the move unit task for the participants and the model. Figure 4 (b) plots the distribution of the participants' fixations and the model's attention moves across the task display during a move unit task. As can be seen in Figure 4 (a),

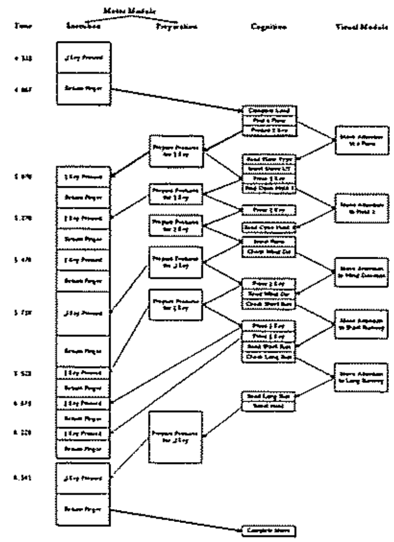

Figure 5. A critical path diagram of an execution of a move unit task. The shadowed boxes indicate the critical path.

the differences in the latencies of functional-level subgoals are minimal between the participants and the model for the move unit task. However, as can be seen in Figure 4 (b), the model differs from the participants in where it is attending to versus where the participants are looking during a move unit task. In particular, the model is attending to the runways and the weather much more than the participants, whereas the participants are fixating on the hold levels much more than the model.

The reason for the discrepancy between the model and the data again lies with ACT-R/PM's current lack of a mechanism to acquire information peripherally. That is, the up-todate information about the conditions of the runways and the weather are required during each unit task execution, because those information are used in the selection of the next unit task. However, because the task provides visual cues when these information change, people can respond to the visual cues peripherally and attend to the runways and the weather only when a change has occurred. The model, on the other hand, must constantly poll (i.e. attend to) these information to see if they have changed.

With a proper mechanism in ACT-R/PM to acquire peripheral information, the total number of attention shifts made by the model to the runways and the weather would be greatly reduced. Therefore, the proportional size of the attention shifts to the hold levels would increase relative to the number of attention shifts to the runways and the weather, thereby making the pattern of attention shifts of the model be consistent with the data.

Critical Path Analysis

A critical path diagram graphically illustrates the critical dependencies in the execution of a model. Figure 5 displays the critical path diagram (John, 1989) of an execution of a move unit task for the model. As can be seen in Figure 5, much of the critical path for a move unit task is on the motor execution component. This implies that when participants become skilled in the KA-ATC task, they are mostly limited by the constraints of the motor system. That is, when people have reached skilled performance in the KA-ATC task, the individual differences in motor speed would greatly predict the performance differences in the KA-ATC task. And indeed, Ackerman (1988) found that the motor ability best predicts skilled performance in the KA-ATC task.

Conclusion

With extended unified theories like ACT-R/PM, we are now able to make detailed models at the level of perceptual-motor processes. Extended unified theories will become particularly important as we extend and scale our cognitive models to real-world, dynamic tasks. This paper represents one attempt to make a detailed model of a dynamic task at the level of perceptual-motor processes. Our efforts in this regard not only provided us with valuable insights into skilled-performance in the KA-ATC task, but also into the current limitations of ACT-R/PM. In particular, the model points to the important role of peripheral vision in performing the KA-ATC task.

Hence, the model highlights the importance of peripheral vision in performing the KA-ATC task. Indeed, peripheral vision is likely to be very important in performing other dynamic tasks as well. This points to a need for such a mechanism in ACT-R/PM, as we begin to model real-world, dynamic tasks.

Acknowledgements

The research reported in this paper was supported by the Office of Naval Research, Cognitive Science Program, under Contract Number N00014-95-1-0223 to John R. Anderson.
Correspondence concerning this article should be addressed to Frank J. Lee, Department of Psychology, 342-C Baker Hall, Carnegie Mellon University, Pittsburgh, PA 15213-3890.

References

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. *Journal of Experimental Psychology: General, 117*, 288-318.

Ackerman, P. L., & Kanfer, R. (1994). *Kanfer-Ackerman air traffic controller task© CD-ROM database, data collection program, and playback program:* Office of Naval Research, Cognitive Science Program.

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

Anderson, J. R., Matessa, M., & Lebiere, C. (1998). The visual interface. In J. R. Anderson & C. Lebiere (Eds.), *The Atomic Components of Thought* (pp. 143-166). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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 Associates, Inc.

Chong, R. S., & Laird, J. E. (1997). Towards learning duel task executive process knowledge using EPIC-Soar. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (p. 107-112). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Douglass, S. (1998). EPAL: Data collection and analysis software for eye-tracking experiments. A computer software, Carnegie Mellon University PA, USA.

John, B. E. (1989). *Contributions to engineering models of human-computer interaction: I and II.* Ph.D. Thesis, Carnegie Mellon University, PA, USA.

John, B. E., & Lallement, Y. (1997). Strategy use while learning to perform the Kanfer-Ackerman Air Traffic Controller© task. In *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society* (p. 337 - 342). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Lee, F. J. (1999). *Does learning of a complex task have to be complex? A study in learning decomposition.* Ph.D. Thesis, Carnegie Mellon University, PA, USA.

Lee, F. J., Anderson, J. R., & Matessa, M. P. (1995). Components of dynamic skill acquisition. In *Proceedings of the Seventeenth Annual Conference of the Cognitive Science Society* (p. 506 – 511). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: I. Basic mechanisms. *Psychological Review, 104(1)*, 3-65.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Reder, L. M., & Schunn, C. D. (1998). Bringing together the psychometric and strategy worlds: Predicting adaptivity in a dynamic task. In D. Gopher & A. Koriat (Eds.), *Cognitive regulation of performance: Interaction of theory and application. Attention and Performance XVII.*

Schunn, C.D., & Reder, L.M. (submitted). Another Source of Individual Differences: Strategy Adaptivity to Changing Rate of Success.