Milliseconds Matter But So Do Learning and Heuristics

Bella Z. Veksler (zafrib@rpi.edu)
Rensselaer Polytechnic Institute
110 8th St., Troy, NY 12180

Abstract

Prior work has shown that the interleaving of perceptual, motor, and cognitive components results in a considerable speedup in the performance of a simple decision making task (Veksler, Gray, & Schoelles, 2007). The current modeling effort conducted using the ACT-R cognitive architecture (Anderson & Lebiere, 1998) is intended to demonstrate how this interleaving might be learned, and how decision-making in this task might take place. The model learns the interleaving and exhibits a speedup in performance similar to that of human participants (RMSE=4.3sec). Furthermore, the model matches human accuracy by using a simple heuristic to make decisions.

Introduction

Previous work has shown that milliseconds matter in understanding human performance (Gray & Boehm-Davis, 2000; Veksler et al., 2007). This millisecond improvement has been shown to occur in a table-based, decision-making task (Lohse & Johnson, 1996) without resorting to changes in higher-order decision-making strategies. Furthermore, exploratory modeling revealed the necessity to focus on the millisecond level considerations in skilled task performance. It was found that an important aspect of the model in mirroring the speedup in performance observed in human participants was the interleaving of cognitive, perceptual, and motor operations. An additional speedup was observed in human data as participants minimized the distance they moved the mouse while interacting with the interface.

Our current modeling effort seeks to extend this by (1) including a learning component to the model whereby the model learns the interleaving and distance minimization on its own, and (2) implementing a higher order strategy to match human accuracy performance.

The Task

The experimental environment used in this research was designed to study and model how information access influences the way in which a decision is made – specifically what information is considered and how it is integrated given the environmental constraints and accessibility of information. In particular, we were interested in whether or not people would take advantage of particular regularities in the environment in order to maximize their score. We hypothesized that this exploitation would occur more when the cost of information acquisition was higher (longer lockout durations).

We used a simple table task (see Figure 1) similar to the one used in a previous study (Veksler et al., 2007) with a few important alterations. The current task environment contained five alternatives (arranged in rows) with a value on each of five attributes (arrayed in columns). In addition, each attribute had an assigned probability value which indicated that attribute's relative importance to the alternative's total score. However, the values in the grid were not visible to the participant and they could only uncover one value at a time. The task environment also allowed us to manipulate the duration of the lockout between a participant selecting a cell in the grid and the value of that cell appearing on the screen, so as to allow us to determine the cognitive and perceptual-motor tradeoffs involved.

In the previous study we conducted in the lab, we manipulated how information was accessed – whether participants could see an entire row, an entire column, or only one cell at a time. In the current study, we instead wanted to explore what particular pieces of information people would gravitate towards given a different cost of exploring the grid – how long they had to wait for information to appear. We hypothesized that the cost of information acquisition would influence not only the exploration of the task environment but also the accuracy of the decisions.

Another important change from the original study, is that we went back to the original decision-making table task and implemented different 'gambles,' composed of various sets of probability values for the attributes, in order to see how they would affect performance (Payne, Bettman, & Johnson, 1988) since that work indicated that the probability landscape of the task influenced the strategies people used to complete the task.

Method

We used a traditional decision-making table task for the study.

Participants

A total of 75 undergraduates (22 females and 53 males) from Rensselaer Polytechnic Institute participated in the study. The average age was 19.21 years (*SD* = 2.05). Students received extra credit for their participation.

Design

There was one between-subjects independent variable of lockout duration with 3 levels. The levels varied the duration of the lockout prior to a value appearing on the screen when a participant clicked on a cell. The three lockouts were 0s (0-lock), 2s (2-lock), and 4s (4-lock). However, for purposes of the models we only focused on the 0-lock condition. There was a within-subject independent variable of gamble type with 4 levels. The gamble types are listed in Table 1. Each gamble type consisted of 5 column (outcome) probabilities that were randomized on each trial within a block of 10 trials. The dispersion of each gamble type refers to the degree to which one of the column probabilities 'dominates' the others. For example, Gamble Type 0 has one column probability of .6, which is significantly greater than any of the other column probabilities. Gamble Type 0 therefore has a higher dispersion value than any of the other gambles since cell values in the column containing a probability of .6 would contribute more to the final value of an alternative (row) as compared to any other columns. The order of the gambles was randomized within each epoch of 40 trials (10 consecutive trials in each block contain the same gamble type). There were two epochs in the study resulting in 80 trials.

Table 1 : Gamble types used in the study. Column probabilities are randomized from trial to trial within a block of 10 of a particular gamble type. The dispersion value is the standard deviation of the 5 probabilities comprising the gamble.

Gamble Type	Column Probabilities	Dispersion
0	.6,.1,.1,.1,.1	.22
1	.4,.3,.1,.1,.1	.14
2	.3,.2,.2,.2,.1	.07
3	.2,.2,.2,.2,.2	0

Materials

The experiment was presented using a computer running Mac OS X on a 17" flat-panel LCD monitor set to 1024x768 resolution. The software used for the experiment was written in LispWorks 5.0. Each trial consisted of a blank grid being presented to participants (Figure 1).

Along the top of the grid were listed the corresponding column probabilities for that column. The alternatives to choose among were the rows in the grid and participants had to click on the radio button to the left of the alternative to make their choice. Each cell in the grid could be uncovered by clicking on it. Once a cell was clicked, any cell clicked prior to the current one would be covered up. Therefore, only one cell value was visible at any given time. Since we found that in our original study, the task was easier for the participants than we originally anticipated, in order to make the current version a bit more difficult, the cell values were randomly selected from the range 11 to 50 rather than being one of 0, 2 or 4.

Procedure

Each participant was run separately. Participants were asked to turn off their cell phones for the duration of the study. After signing informed consent forms and going through the instructions on how to do the task, each participant completed 80 decision-making trials. These were broken down into blocks of 10 and each block of 10 had one of the 4 gamble types. Participants were instructed to choose the alternative (row) that had the highest weighted summed value. Specifically, the expected value of any given alternative can be calculated by:

$$EV(alt_j) = \sum_{i=1}^{5} p_i v_{ji}$$

p: outcome (column) probability in column i
v: cell value of cell in row j and column i

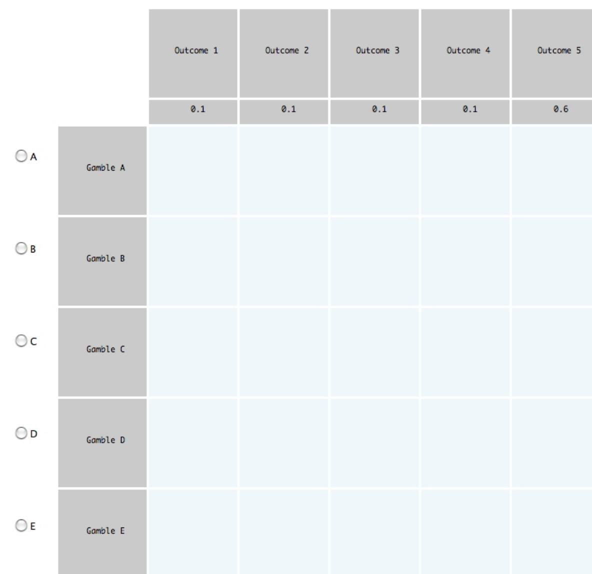

Figure 1: Task Environment

The reward given for each trial was the ratio of the alternative chosen by the participant compared to the best alternative's expected value. Therefore, if the participant chose the best alternative they received a reward of 100 points, if the next best alternative (and its ratio to the best was 98) then they would receive 98 points.

Participants were given feedback on their score after each trial, along with how long they spent on the trial and how many cells they uncovered. At the end of a block of trials they were given feedback on their average score for that block. At the end of each epoch they were given feedback on the average score over the 40 trials.

Results

Several participants had to be excluded from the analysis due to software malfunction. Consequently, only data from 58 participants (16 females and 42 males) was used for the analysis, 20 participants in the 0-lock condition, 19 in the 2-lock and 19 in the 4-lock. However, it should be noted that the current modeling work only addresses the 0-lock condition of this study. Future work will also incorporate the other conditions.

Accuracy

A 4x3 repeated measures ANOVA on the effects of lockout and gamble type on average accuracy over 80 trials was conducted. The repeated variable was gamble type. There was not a significant gamble*lockout interaction, $F(6, 165) = 1.11$, $p = 0.358$. There was a significant main effect of gamble type, $F(3, 165) = 62.2$, $p < 0.001$. There was also a significant main effect of lockout, $F(2, 55) = 6.87$, $p < .01$. Figure 2 illustrates the trends in accuracy across the four gamble types with respect to the lockout condition.

There was a significant linear trend, $F(1, 55) = 179.6$, $p < .01$, $\omega = .46$, indicating that as the dispersion of the gambles decreased, average score increased. Post-hoc tests revealed significant differences between 0-lock and 4-lock conditions, with a mean difference of 3.28, $p < .01$.

These results indicate that participants in longer lockouts had on average less accurate choices and that accuracy was worse for gambles that had more 'dominating' probability columns.

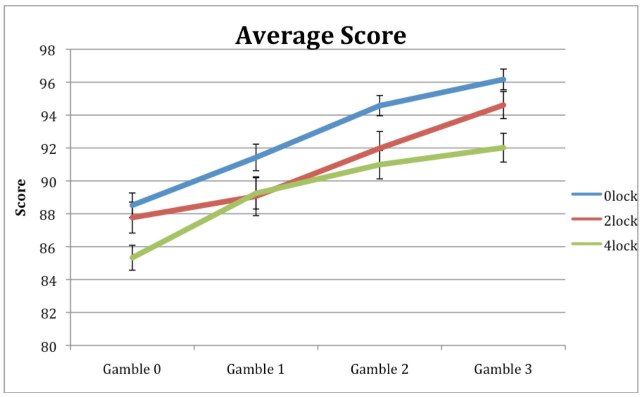

Figure 2: Average Accuracy across Gambles and Lockout Conditions. Error bars are standard error.

Duration of Trial

A 8x3 repeated measures ANOVA was conducted on the effects of lockout and block on how long cell values appeared on the screen. The repeated variable was block number. There was not a significant block*lockout interaction, $F(5.7, 156.66) = 2.07$, $p = 0.06$. There was a main effect of block, $F(2.85, 156.66) = 25.41$, $p < 0.01$. There was also a significant main effect of lockout, $F(2, 55) = 11.94$, $p < .01$. Figure 3 illustrates the trends in average trial duration. Of note here is that there is a significant speedup over the course of the study, in all of the conditions.

Location of Cell Clicks

In order to better understand the strategies people were using to do the task, we looked at which cells participants tended to uncover. In the previous study (Veksler et al., 2007), we found that when given the opportunity to view values by rows vs. by columns, participants chose to check cell values within a row before transitioning to the next row, rather than clicking consecutive cells within a column. We subjected the data of the 0-lock group from the current study to the same analysis. We examined the percent of cell clicks that were either on two consecutive cells in a row or in a column (henceforth referred to as cell transitions). We found that about twice as many cell transitions occurred within a row rather than within a column (Figure 4).

A paired sample t-test revealed a significant difference between the percent of cell transitions within a row ($M = .59$, $SE = .04$) versus within a column ($M = .29$, $SE = .04$), $t(19) = 3.96$, $p < .001$. This suggests that people tended to use a by-row strategy of evaluating alternatives rather than focusing on the columns and our current modeling effort reflects this strategy as well.

We were also interested in whether participants tended to consider the probability values assigned to the columns in their decision making process. In particular, we hypothesized (and previous work by Payne et. al. has shown) that gambles that had higher dispersion values should have more cells uncovered containing the higher probability columns. For the sake of brevity, our findings were that there was not a significant difference between the percent of cells participants clicked in the different probability columns as compared to what would be expected by chance.

We also hypothesized that cells in the higher probability columns would be uncovered earlier in the trial rather than later. However, we found that although there was a considerable bias toward checking grid values starting at the

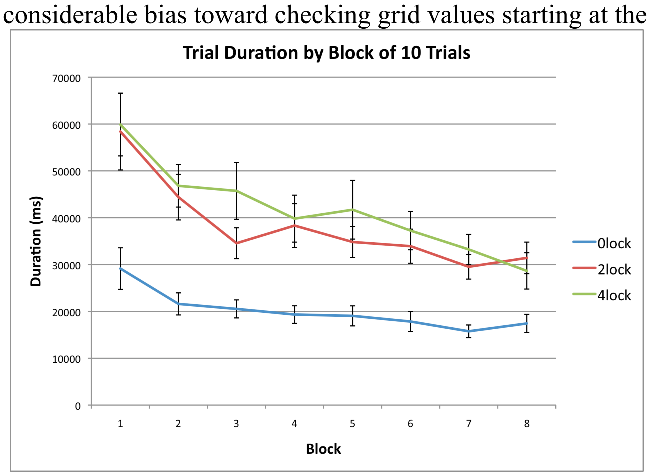

Figure 3: Average duration of trial by block of 10 trials. Error bars are standard error.

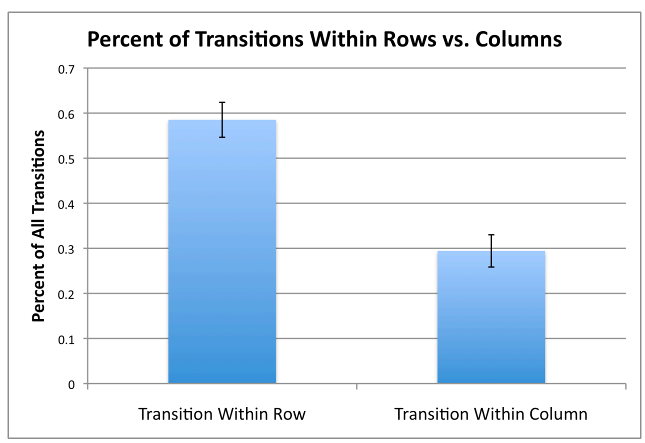

Figure 4: Percent of cell click transitions occurring within a row versus within a column. Error bars are standard error.

top row and moving down (average first click on top row = 1.07, average first click on bottom row = 12.08) and a bias toward checking cells in the left hand columns first (left column = 5.02, right column = 8.7), there was not a significant bias toward checking higher probability columns first.

The Model

To model human performance on this task, we used the ACT-R cognitive architecture (Anderson et al., 2004). ACT-R is a modularized production system with a subsymbolic memory module. It has visual and motor modules to embed it in the task environment. It also has declarative memory and a procedural module. In addition, it has imaginal and goal buffers to store its working memory and goal chunks, respectively. Thus, it serves as a good framework to model human performance on this simple table task.

The current modeling work combined the static models of previous modeling work (Veksler et al., 2007), to demonstrate the learning component in order to fit human data on the task. Furthermore, whereas the previous modeling effort was more concerned with the speed of the interactive routines, the current model also attempts to reproduce accuracy.

The structure of the current model is similar to that of the previous models and is briefly described here. There are roughly four components to the model: switching between alternatives, moving through the cell values within an alternative, comparing the current alternative's value to the best so far, and answering. Figure 5 illustrates the flow of the model and the various productions involved. There are two important changes from the previous models (Veksler et al., 2007) to the current model. The first is the introduction of two sets of competing productions intended to produce a learning effect in the model. The second is the change in strategy implemented by the model to complete the task. We will address each of these important changes in turn.

Competing Productions – Learning Speedup

In matching trial duration of the human data, we implemented two sets of competing productions intended to demonstrate the speedup in performance.

The first two productions that compete occur in the "Switching Between Alternatives" part of the model. As per the previous modeling effort, we found that human participants initially clicked on cells in a left to right fashion whereas later they alternated the direction depending on their ending position in a given row. We thus incorporated this alternating behavior into the model thereby decreasing the distance the mouse had to move when a new alternative was encountered. Since move-mouse execution time in ACT-R is closely related to the distance that the mouse must move, as per Fitts' Law (Fitts, 1954; MacKenzie, 1992), this feature allowed the model to transition faster between alternatives (about 900ms faster over the course of the trial). The two competing productions 'change-row l->r' vs.

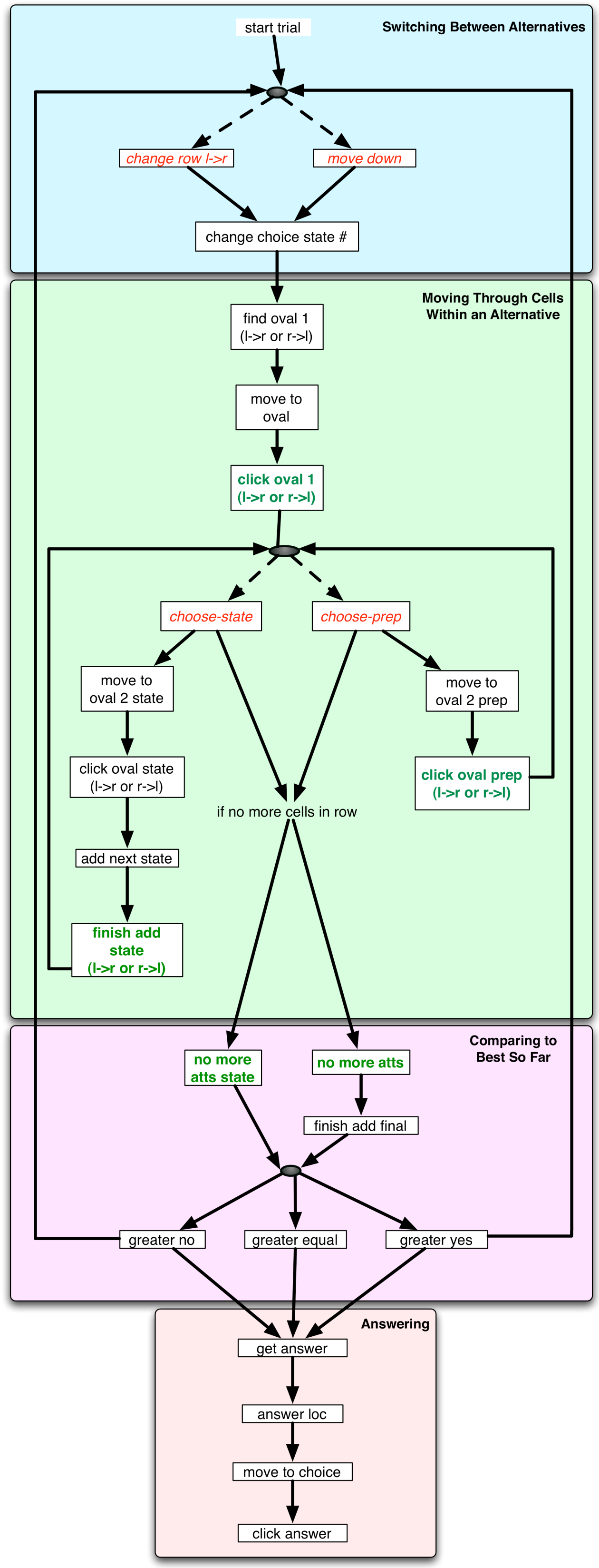

Figure 5: Schematic of the Model. Dashed lines indicate competing productions. Productions in green propagate a reward. Productions in red are competing

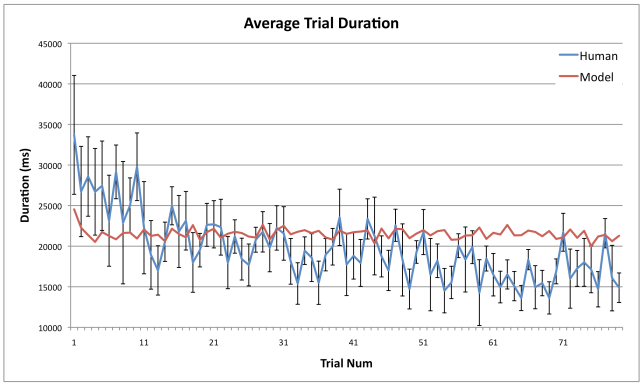

Figure 6: Average trial duration comparison between model and human data. Error bars are standard error.

'move down' are the two types of transitions that we noticed in our human data. Initially the utility of the 'change-row l->r' production is considerably greater than the 'move-down' production, however, the model quickly learns the greater utility of choosing to move down to the next row rather than always resorting to reading the cell values left->right.

The second set of productions that compete occurs in the "Moving Through Cells Within an Alternative" part of the model. Again, as per our previous modeling effort, we noticed that a considerable speedup in performance could be attained by having the model interleave cognitive, perceptual, and motor components (Veksler et al., 2007). The two competing productions are 'choose-state' and 'choose-prep.' The productions following the 'choose state' production all have no interleaving of the perceptual-motor-cognitive components whereas the productions following the 'choose-prep' production do include all the interleaving as described in previous work, and as can be seen in Figure 5, comprise half as many productions.

ACT-R uses a reinforcement learning mechanism for updating production utilities and is based on the amount of reward and time since the production fired that the reward has been triggered as well as a noise parameter. The utility of a production *i* at time *n* is defined by the equation (Bothell, 2004):

$$U_i(n) = U_i(n-1) + \alpha[R_i(n) - U_i(n-1)]$$

α is learning rate (set to .2)

$U_i(0)$ is set to 1000 for 'choose state' and 1 for 'choose prep'

$R_i(n)$ is the effective reward given to production *i* at time *n* calculated by subtracting the reward at time *n* minus the time since production *i* was selected

In order to even the playing field, in all cases the same amount of reward is triggered by the rewarding production (in this case we used a reward of 1). However, based on the current model's competing productions, it turns out that the major factor influencing how much reward each of the competing productions receives (and thereby alters its utility) is the time since the competing production fired compared to the reward production. The average difference between how long this interval was for 'change-row l->r' vs. 'move down' is 85ms. The average difference between how long this interval was for 'choose-state' vs. 'choose-prep' is 471ms. Over the course of the 80 trials, the model quickly learns the higher utility of using the 'move down' and 'choose prep' productions.

Figure 6 illustrates the average trial duration for both human and model data, which is a direct result of which of the competing productions are selected during a particular trial. Qualitatively, there is a learning curve for both humans and the model over the course of the first few trials, *RMSE* = 4.35s and the correlation coefficient is .21. The low level analysis of the time it takes both the model and the human participants to transition between consecutive cells in the grid indicates similar trends, *RMSE* = 131.74ms and the correlation coefficient is 0.28. Past work has addressed this low level analysis and for brevity only the fit is mentioned here (Veksler et al., 2007). Future work will need to address how to account for the remainder of the speedup seen in human data, perhaps as strategy shifts come into play later during the course of the experiment.

Model's Strategy – Accuracy Matching

The model just described was also outfitted with a simple heuristic in order to match human accuracy on the task. The strategy change that we implemented had to do with our analysis of cell clicks in the human data and the current task environment's setup. In particular, since we no longer had easy values in the cells of the grid, computing the normative value of an alternative is much more difficult than in our original task. Instead, given our human data analysis and how quickly participants were transitioning between cells in the grid, we suspected that rather than multiplying out the values and probabilities and summing these across the alternative, our participants were using a simpler heuristic to determine the best alternative.

This heuristic strategy was implemented in the model whereby as the model uncovered cell values, it simply kept a count in its imaginal buffer as to the number of cells in a particular row whose values exceeded some predetermined threshold value. Thus, rather than doing any sort of computation per se, the model was merely keeping count. At the end of a trial, the choice the model made was based on the alternative that it found to have the most cells above a threshold. If there were ties among alternatives, the more recent alternative looked at was chosen.

The implementation of this strategy also led to an important consideration – where to place the threshold. We explored the threshold parameter space in closed form to determine which threshold resulted in the best fit to human accuracy data. The procedure used is described below.

Threshold Consideration

A closed form model of the threshold parameter was developed to explore the model's accuracy given one of 35 threshold values (15 to 49). At first, 24 random 80-trial stimuli were used and run through each of the 35 threshold values and it was determined that a threshold value of 40 provided the best fit to average human performance, *RMSE* = 0.61. We then took all of the stimuli from the human participants (actual trials participants saw) and ran those through the model using the threshold of 40. Figure 7 depicts the fit of the model with a threshold of 40 to human data.

A 2x4 repeated measures ANOVA was conducted to compare human and model accuracy (type) with the repeated measure being gamble. There was not a significant gamble*type interaction, $F(3, 151) = 1.16$, $p = 0.33$. There was a significant main effect of gamble, $F(3, 151) = 143.65$, $p < 0.001$. There was not a significant main effect of type, $F(1, 151) = 0.08$, $p = 0.78$.

This analysis indicates that there was not a significant difference between human and model accuracy across the 4 gamble types. However, there was a significant difference between the gambles for both humans and the model.

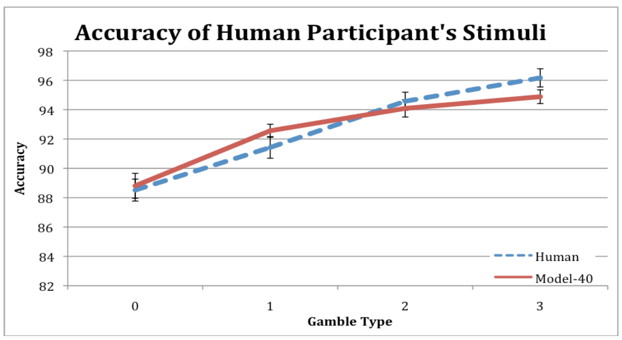

Figure 7: Accuracy comparison of model with threshold 40 across all 80 trials of human participant's stimuli. Error bars are standard error.

Conclusions

The current modeling work had a twofold purpose. The first was to demonstrate that the model could learn the cognitive, perceptual, motor interleaving resulting in the speedup in performance shown in previous work. The second was to implement a decision-making strategy that human participants most likely utilized in order to do the task.

Given the human data collected from a study of a decision-making table task, we found accuracy differences dependent on the constraints of the task environment (both lockout durations and types of gambles used). We also found that over the course of the 80 trials, participants completed trials considerably faster. The current model also completes the trials faster over the course of the task.

Furthermore, a more rigorous analysis of the human data indicated some biases in the way participants interacted with the task environment and we have implemented these biases in the strategy the model uses to complete the task. Namely, the model goes through the grid of cells in a top-down manner, and begins with the left-most column in the first row that it uncovers. In addition, the lack of a bias to click on the higher probability columns and the fact that gambles with higher dispersion values also had lower average scores, indicates that human participants tended to disregard the probability data, at least as far as the 0 second lockout group was concerned, and our model did as well. Future work will need to address how to reconcile this result with previous results of Payne et. al. (1988) in which it was found that probabilities played a role in decision strategies.

Future work will also incorporate the data we have from the other two conditions of the study as it relates both to strategy selection and timing. We also plan to further explore the factors influencing how quickly the model can perform the task as it seems human participants are nevertheless faster.

Acknowledgements

Thanks to Wayne Gray, Mike Schoelles, and Vladislav Veksler for many useful discussions regarding this ongoing project. The work was supported, in part, by grant N000140710033 to Wayne Gray from the Office of Naval Research, Dr. Ray Perez, Project Officer.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglas, S., Lebiere, C., & Quin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036-1060.

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

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47*(6), 381-391.

Fu, W.-T. (2007). A rational-ecological approach to the exploration/exploitation trade-offs: Bounded rationality and suboptimal performance. In W. D. Gray (Ed.), *Integrated models of cognitive systems*. (pp. 165-179). New York, NY: Oxford University Press.

Gray, W. D., & Boehm-Davis, D. A. (2000). Milliseconds matter: An introduction to microstrategies and to their use in describing and predicting interactive behavior. *Journal of Experimental Psychology: Applied, 6*(4), 322-335.

Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes, 68*(1), 28-43.

MacKenzie, I. S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction, 7*(1), 91-139.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*(3), 534-552.

Veksler, B. Z., Gray, W. D., & Schoelles, M. J. (2007). *From 1000ms to 650ms: Why Interleaving, Soft Constraints, and Milliseconds Matter.* Paper presented at the 8th International Conference on Cognitive Modeling, Ann Arbor, MI.