In: B. Kennedy, D. Reitter, R. St. Amant (Eds.), *Proceedings of the 22nd Annual Conference on Behavior Representation in Modeling and Simulation*. Ottawa, Canada: BRIMS Society, 2013.

An ACT-R Model of Sensemaking in Geospatial Intelligence Tasks

Jaehyon Paik, Peter Pirolli
Palo Alto Research Center
3333 Coyote Hill Rd.
Palo Alto, CA 94304
jpaik@parc.com, pirolli@parc.com

Wei Dong
Cornell University
301 College Avenue
Ithaca, NY, 14850
wd96@cornell.edu

Christian Lebiere, Robert Thomson
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15218
cl@cmu.edu, thomsonr@andrew.cmu.edu

Keyword:
Instance Based Learning, Reinforcement Learning, Sensemaking, ACT-R, Anchoring Biases

ABSTRACT*: We developed an ACT-R model of sensemaking in geospatial intelligence tasks based on two widely used learning processes in ACT-R: instance-based learning and reinforcement learning. This map-based task requires users to select (make visible) layers that visualize different types of intelligence, and to revise probability estimates about which groups might commit a future attack. The model (a) evaluates the gains to be made by selecting layers during the simulation, (b) selects layers based on the evaluation of all layers, and (c) adjusts probability estimates of the threats posed by all groups based on new evidence. The model exhibits layer-selection patterns that are comparable to participants (N = 45) studied on this task and both model and people deviate from a rational model based on greedy maximization of expected information gain. The model also exhibits an anchoring bias in updating belief probabilities based on revealed evidence, which corresponds to the average participant.*

1. Introduction

Sensemaking (Klein, Moon, & Hoffman, 2006a, 2006b; Pirolli & Card, 2005; Russell, Stefik, Pirolli, & Card, 1993) is a concept that has been used to define a set of activities and tasks in which there is an active seeking and processing of information to achieve understanding about some state of affairs in the world. Various kinds of complex tasks in intelligence analysis and situation awareness have been frequently used as examples of sensemaking (Klein et al., 2006a, 2006b; Pirolli & Card, 2005). According to Pirolli and Card (2005), the overall process of sensemaking can be organized into two major loops of activities, a foraging loop and a sensemaking loop. The foraging loop involves seeking, searching and filtering information, and reading and extracting information into representation called a schema. The sensemaking loop involves iterative development of schemas to make best fit with the evidence.

A recent study (Thomson, Lebiere, Rutledge-Taylor, Stazewski, & Anderson, 2012) showed that the ACT-R cognitive architecture can capture several basic components of sensemaking theory by presenting how the various cognitive mechanisms of an ACT-R model applied to a geospatial intelligence task can be used to in sensemaking theory. Here we present an ACT-R model of that geospatial intelligence task that requires foraging for information and compare it to human performance data collected in a controlled study.

2. The Geospatial Intelligence Task

The IARPA ICArUS program provides a series of six challenge tasks to drive the development of integrated neurocognitive models of sensemaking. Participants performed Tasks 1 through 6 with resting period between the tasks. The final task, Task 6 (Figure 2.1 (d)) requires reasoning based on a set of rules (PROBS rules) that relate observed evidence to the likelihood of attack by four different groups.

A layered geospatial map is presented on a computer screen with different layers presenting different forms of intelligence (INTs). There are five types of INTs (see Table 2.1) that include HUMINT (human intelligence),

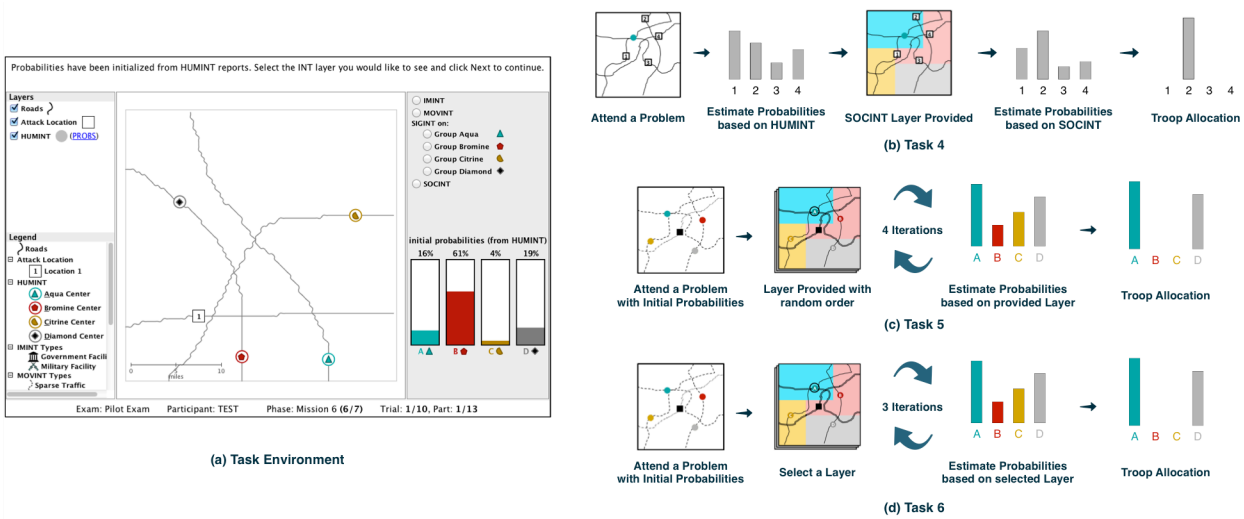

Figure 2.1 The task environment (a) and a series of geospatial intelligence tasks (b, c, and d). Task 4 (b) begins with estimating probability of attack for four different locations based on distances between the attack location and the insurgent group. A SOCINT layer was provided, then participants were asked to update probabilities of four different locations again base on SOCINT rule. In Task 5 (c) and Task 6 (d), participants initially were provided HUMINT information. Based on the provided/selected layer, participants were asked to update probability of each group. After end of each trial, participants were asked to allocate resources (troops) base on their probability estimation.

IMINT (image intelligence), MOVINT (movement intelligence), SIGINT (signal intelligence), and SOCINT (socio-cultural intelligence).

Participants performed two similar tasks (Task 4 and 5, see Figure 2.1) before the Task 6. Task 4 begins with four different possible attack locations and a given insurgent group center. Participants were asked to report the probability of attack at each location based on the distance between the insurgent group center and each location. After reporting probabilities, participants were given a SOCINT layer that indicates the group in whose region the attack had happened. Participants were asked to report probabilities again based on the SOCINT evidence and the SOCINT PROBS rule for reasoning about that evidence (Table 2.1).

Tasks 5 and 6 are very similar. Both tasks begin with an attack location and four possible insurgent groups with their attack probabilities (HUMINT). In Task 5, the remaining four types of INTs were provided, one at a time in a random order. In Task 6, participants were asked to choose a particular INT (layer), up to a total of three INT layers, one at a time, in any order. Each INT layer provides unique outcome. Specifically, IMINT can reveal whether an attack happened on a government or military building, MOVINT provides outcomes on whether an attack occurred in dense or sparse traffic, SIGINT indicates electronic "chatter" or "silence" by different groups, and SOCINT provides the group in whose region the attack happened. At each stage, the selection of a particular INT provides outcomes that can be used to update probability distributions over the hypotheses about the responsibility of the four different groups in producing the given attack. The rules specifying how outcomes ought to update these probabilities are given as the PROBS rules in Table 2.1. After the last stage of both tasks, participants were asked to allocate resources (troops) to prevent further attacks and the allocation score was provided to participants based on their allocation for each group and the ground truth (e.g., if the ground truth is A, the allocation is 40-30-20-10 for each group, then the allocation score is 40).

Table 2.1. Probabilistic rules provided to user for inferring beliefs about group attack likelihoods.

INTS	PROBS
HUMINT	If a group attacks then the likelihood is a normal (Gaussian) function of distance along road(s) from the group's center.
IMINT	If A or B attack then the attack is four times as likely to occur on a *Government* vs. *Military* building. If C or D attack then vice versa.
MOVINT	If A or C attack then the attack is four times as likely to occur in *dense* vs. *sparse* traffic. If B or D attack then vice versa.
SIGINT	If SIGINT on a group reports *chatter*, then attack by that group is seven times as likely as attack by each other group If SIGINT on a group reports *silence*, then attack by that group is one-third as likely as attack by each other group.
SOCINT	If a group attacks then that group is twice as likely to attack in its own vs. other region.

3. Analysis of Human Data

Forty-five participants (MITRE Technical Report, In Progress) performed the Tasks 1 through 6 with resting periods between tasks. We analyzed participants' layer selection sequences in Task 6 first. Figure 3.1 shows the observed layer selection sequences for Task 6, which indicates that about one third of the all sequences follow the IMINT-MOVINT-SIGINT sequence (the vertical order of layers appearing in the Graphical User Interface (GUI), see Figure 2.1), another third selected SOCINT for the first choice and selected ANYINT (any layer) for the rest of their choices. The remaining layer selection sequences appear to be random selections.

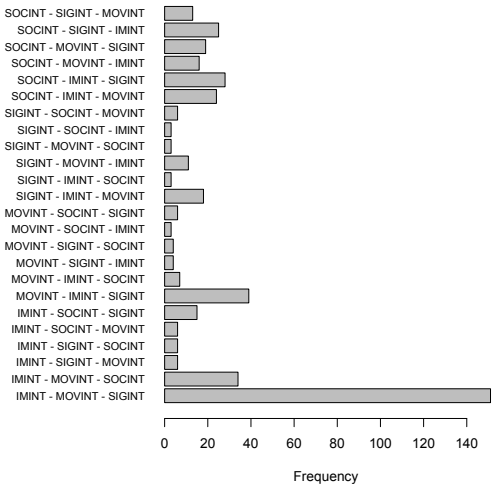

Figure 3.1 Layer selection sequences in human data on Task 6.

We hypothesize that some participants selected the IMINT-MOVINT-SIGINT sequence more often because they were influenced by the experimental GUI. As we can see in Figure 2.1, the IMINT-MOVINT-SIGINT sequence is the vertical order of layers presented on the GUI, so some of participants just followed that sequence without other considerations, such as information gain.

We also assume that participants' previous experiences made them have some preference for a particular layer. That is, if the participants had some positive experience for specific layers in Tasks 4 and 5, they might choose those preferred layers rather than the others in Task 6. We analyzed the outcomes of SOCINT layer selections in Tasks 4 and 5, because SOCINT-first choices were observed frequently in the human data, even though it provides the least expected information gain among all layers (based on the expected change in the entropy of the probabilities assigned to groups if one follows the rules in Table 2.1). Figure 3.2 shows the frequency that hypotheses about group responsibility are assigned the highest probability immediately after initial HUMINT (distance estimation) evidence in Task 4. Figure 3.2 shows that in more than 85% of the cases the hypothesis of group D responsibility has the highest probability. When the SOCINT layer was presented to participants, the outcome was always region "D", supporting the highest probability group.

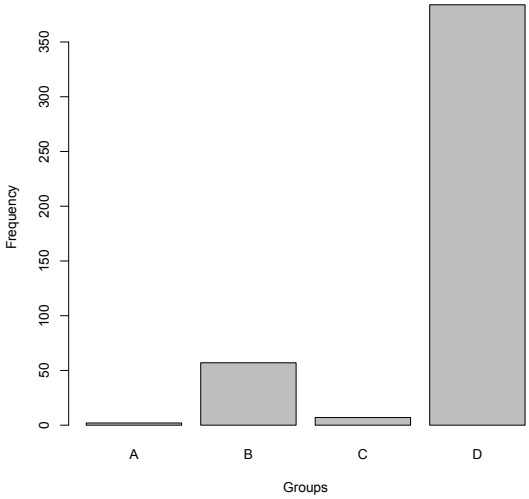

Figure 3.2 Frequency that a group was the highest probability group after HUMINT (distance estimation) in Task 4.

Figure 3.3 plots Task 5 data showing how SOCINT evidence supports the group-responsibility hypothesis. The x-axis in Figure 3.3 plots the SOCINT support according to the rank probability of the hypothesis at the time of SOCINT presentation.

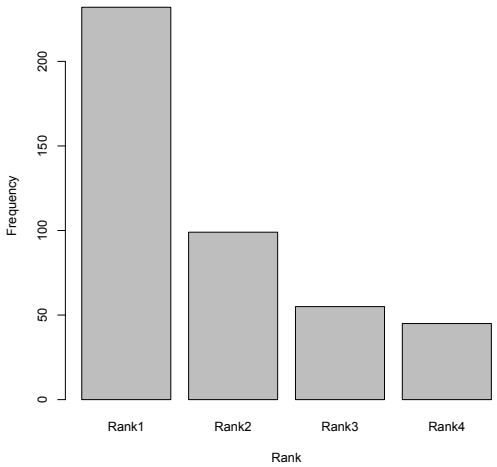

Figure 3.3 Frequency of the SOCINT outcome supports the highest to the lowest probability by rank order.

Figure 3.3 shows that the outcome of the SOCINT layer supports the highest probability group more than 53% of the time and the second highest probability group more than 23%. These results suggest that participants had frequent positive experiences of the SOCINT layer in Task 5, which may have made them choose a SOCINT layer more frequently in the layer selection of Task 6.

We also investigated the reason that participants rarely chose the SIGINT layer as their first choice, because the SIGINT layer gives the highest expected information gain among all layers. One possible reason is that the SIGINT layer appears to involve a high mental calculation cost, because participants needed to decide on a particular group first, and then consider the possible outcomes of the layer selection. We also assumed that experience from the previous tasks (Tasks 4 and 5) and the ongoing task (Task 6) might make participants not to choose the SIGINT layer as their first choice. We investigated the actual information gain (as measured by change in entropy of the probabilities assigned to groups) at the end of each trial with respect to the first layer choice. Figure 3.4 shows that participants get the most information gain when SOCINT was presented as the first layer in Task 5. SIGINT ranks as second in Task 5, but the worst in Task 6. Therefore, if participants were learning from experience, there is sufficient evidence suggesting that selecting the SIGINT layer, as the first choice, might not be an optimal choice. So, SIGINT is not necessary for seeking the highest expected information gain when making the first layer selection, because it cannot guarantee the optimal end results, and this might explain why participants did not frequently use the SIGINT-first strategy in Task 6.

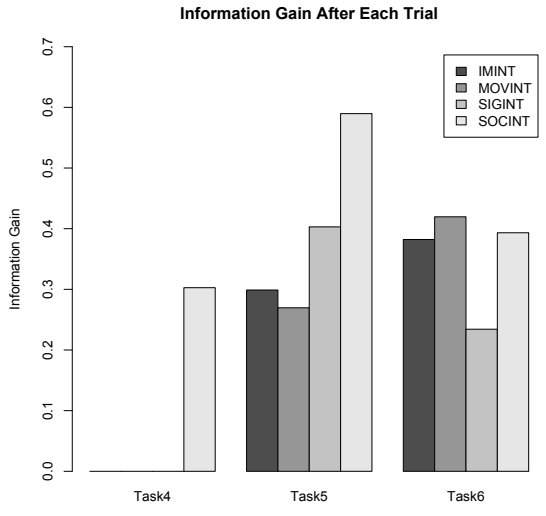

Figure 3.4 Information gain at the end of the trial with respect to layers and tasks

4. ACT-R Architecture

ACT-R (Anderson, Bothell, Byrne, Douglass, Lebiere, & Qin, 2004; Anderson & Lebiere, 1998) is a cognitive architecture. It includes a declarative memory module that stores and retrieves information, a procedural module that coordinates the flow of information, and perceptual-motor modules that enable the model to interact with the external environment. Each module has a buffer, which contains a chunk, as interface to the procedural module and the rest of the architecture to receive commands and requests and return results. Each chunk has an associated numerical value, called activation, that reflects its expected degree of usefulness at any particular moment. When a retrieval request is made, the most active matching chunk is retrieved. Partial matching is a mechanism that allows retrieving a chunk that does not perfectly match a retrieval request. Based on a combination of the activation strength and a similarity score, the best chunk is selected. Blending (Lebiere, 1999) is a memory retrieval mechanism that allows retrieval of an aggregation of all possible chunks in declarative memory, weighted by their probabilities of retrieval reflecting their activation strength and similarity.

The information flow in ACT-R is controlled by a production system. Each production consists of if-then condition-action pair. Conditions are criteria against matched chunks in buffers (e.g., goal, retrieved memory chunk, visual object chunk), and actions make changes to the contents of buffers that trigger operation in the associated modules. The production with the highest utility is selected among possible productions that match the current conditions.

5. ACT-R Model

Our previous ACT-R model (Paik, Pirolli, Lebiere, & Rutledge-Taylor, 2012) was based on assumptions about the behaviors of experts who invariably rely on vast amount of declarative memory experience and well-practiced cognitive skill (Klein, 1999). However, our observed participants had no chance to have a vast amount of declarative knowledge experience during the experimental tasks. So we are proposing an alternative way to model the layer selection process.

5.1 Layer selection process

We considered four cognitive processes to develop an ACT-R model.

- *Difference reduction heuristics.* We assume that participants used a heuristic such as hill-climbing to evaluate layers rather than maximization of expected information gain, because an average person is not able to compute the expected information gain for all layers. Hill-climbing analysis enables participants to

focus on achieving states that are closer to an ideal goal state. This analysis just requires a simple evaluation of difference between the current state and the perfect (goal) state.

- *Instance-Based Learning*. We hypothesize that participants might rely on direct recognition or recall of relevant experience from declarative memory or, failing that, heuristically interpret and deliberate through the rules and evidence provided in the tasks. This compute-vs-retrieve process is a design pattern that typically structures ACT-R models. The notion that learners have a general-purpose mechanism whereby situation-action-outcome-utility observations are stored and retrieved as chunks in ACT-R declarative memory derives from instance-based learning theory (*IBLT*, Gonzalez, Lerch, & Lebiere, 2003; Reitter, 2010). Following the actual task experiences of participants, the model had opportunities to acquire 20 instances for the SOCINT layer during Tasks 4 and 5, and 10 instances for the other layers during Task 5. There is also some number of instances from their selections during Task 6. Those instances were stored into declarative memory and were used to simulated look-ahead search in Task 6.
- *Reinforcement Learning.* During Tasks 4-6, participants were asked to update the probability distribution based on the outcome (information revealed) of each layer. Some of the layers and outcomes might support participants' hypothesis, but some of them might not. Reinforcement learning was employed in the ACT-R model to reinforce or punish layer selection productions based on these experiences. This reinforcement learning adjusts the preference order for layer selection.
- *Cost-satisfaction-driven layer selection.* The SIGINT and SOCINT layers require more calculation cost than IMINT and MOVINT layers when computing outcomes. We assume participants might consider the cost-satisfaction factor when exploring a layer selection. Our model incorporates a cost estimate when considering look-ahead search.

Our ACT-R model performed Instance-Based Learning through Tasks 4 to 6, storing declarative chunks that capture current <situation, actions, outcome, utility> experiences, specifically, prior-probabilities, layer-choice, layer-outcome, and layer-utility. For the layer utility, our ACT-R model explored some plausible difference reduction heuristics in a memory-based move evaluation framework. The following equation shows that the goal is to achieve certainty on one of the hypotheses, and distances from the goal of certainty for each hypothesis *i* are captured by 1 - p_i and each distance is weighted by the current probability p_i.

$$\sum_{i \in hypotheses} p_i(1 - p_i)$$

For instance, if the model encountered a situation [.4 .2 .2 .2] as prior (meaning: probability of Group A attack = .4, Group B attack = .2, etc.), IMINT as a layer selection, "government building" as an outcome, and the model updated posterior probability distribution [.5 .3 .1 .1], then the stored chunk is

(exp1
isa layer-choice
prior-a 0.4
prior-b 0.2
prior-c 0.2
prior-d 0.2
layer IMINT
outcomes government
utility 0.64)

Our ACT-R model also performed reinforcement learning through Tasks 4 to 6. After updating the probability distribution over hypotheses about group attacks, based on a layer and its outcomes, the model evaluates whether the model gains information or loses information by comparing the entropy of the prior distribution (prior to selecting a layer) to the posterior distribution (after updating hypotheses). If the model gains information, the production for selecting the current layer receives some reward; if it loses information, the production receives some punishment. This reinforcement learning enables the model to develop a preference order list for all layers, and the preference order of layers was used to determine which layer should be explored first in layer selection process.

In Task 6, the prior probability distribution based on HUMINT (distances between the attack location and each group) is provided by the environment, and the model has a preference ordering acquired from experience on Tasks 4 and 5. Given those priors and a preferred layer, the ACT-R model searches for a similar chunk that has a similar prior and layer-choice from its declarative memory using ACT-R's partial matching mechanism. If the model retrieves a similar chunk, the model relies on blending to retrieve the utility of the current layer. If a similar chunk does not exist, the model needs to decide whether to compute utility or not based on the calculation cost of the current layer. If the model decides to compute, it calculates the utility (weighted distance) of the current layer, creates a chunk (prior-layer-outcome-utility), and adds the chunk into its declarative memory. If the model decides not to compute, it explores the next preferred layer that is in the preference list, and follows the same procedure.

After the model obtains the utility of the current layer, the model evaluates it by comparison to the average utility of all layers to determine whether the utility of the current layer is acceptable or not. If the utility of the current layer is better (smaller weighted distance to certainty) than the average utility, the model is satisfied with the current layer and selects the layer as its choice, if not, the model explores the next preferred layer and follows the same procedure. After the model selects a layer, it creates a chunk and stores it into declarative memory as its previous experience. The model runs this layer selection and probability adjustment process three times to select three different INT layers. After selecting INTs, the model allocates troops based on the final probability distribution. Figure 5.1 shows the probability distribution of layer selection sequences for our ACT-R model, human data, and a rational model based on local maximization of expected information gain.

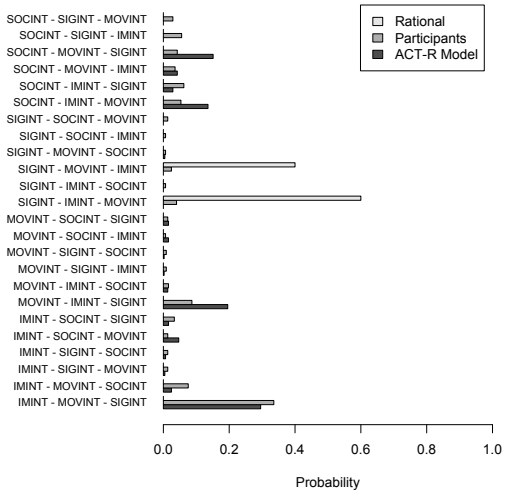

Figure 5.1 Layer selection sequences from the ACT-R model, human data, and rational.

It appears that the ACT-R model focuses more on particular layer sequences than participants. However, participants had more preference for the GUI order (IMINT-MOVINT-SIGINT) than our ACT-R model. To measure the similarity of the probability distribution of layer selection sequences between ACT-R and human data, we measured the Jensen-Shannon Divergence (JSD) between the two distributions. The divergence between the two distributions is .34 (the range of JSD is 0 to 2, 0 means the two distributions are the same), indicating our ACT-R model predicts the human data well.

5.2 Hypothesis probability adjustment

Our model of probability adjustment after receiving new evidence from each INT layer selection is based on a model of cognitive arithmetic (Lebiere, 1999). The cognitive arithmetic model used the retrieval of arithmetic facts to generate estimates of answers rather than explicit computations. The cognitive arithmetic model uses partial matching to retrieve facts related to the problem, and the blending mechanism merges retrieved chunks to get an aggregate estimated answer. The cognitive arithmetic model matched a number of characteristics of the distribution of errors in elementary school children, such as table and non-table errors, error gradients around the correct answer, higher correct percentage for tie problems, and, most relevant here, a skew toward underestimating answers, as is common in anchoring and adjustment processes.

We leveraged the cognitive arithmetic model for our geospatial intelligence model to account for how the PROBS rules (from Table 2.1) are interpreted and applied based on the recent studies (Lebiere, Pirolli, Thomson, Paik, Rutledge-Taylor, Stazewski, & Anderson, Submitted; Rutledge-Taylor, Lebiere, Thomson, Stazewski, & Anderson, 2012) . Initially, our ACT-R model has only five facts that are derived from the instructions provided participants about the PROBS rules (presented graphically during the experiment). Those rule instructions assume cases in which the prior probability distribution over group hypotheses is flat [.25, .25, .25, .25], and present the posteriors for all the outcomes of all the INT layers (e.g., the posterior of IMINT-Government is [.4 .4 .1 .1] when the prior is a uniform distribution). These rules are represented with triplets: an initial probability, an adjustment factor, and the resulting probability. Through Tasks 4 to 6, our model tries to blend over the initial probabilities and the adjustment factor, retrieves the relevant chunks as its posterior, and stores the retrieved chunk into declarative memory if similar chunks (with similar prior) exist. If similar chunks do not exist, the model computes the actual posterior and stores it into declarative memory, then blends the prior with the adjustment factor. Our model computes and stores a lot during the earlier trials, however, it relies more on blending to get the posterior in later trials. When provided with ratio similarities (Lebiere et al., Submitted) between probabilities and factors, the primary effect is an underestimation of the adjusted probability for most of the probability range. This produces a kind of *anchoring bias* as the probability adjustments tend to be closer to the initial prior than what is predicted by normative Bayesian updating.

Figure 5.2 shows the results of probability adjustment for the IMINT and MOVINT layers with respect to participants, ACT-R model, and the Bayesian rational adjustment. The x-axis is a prior probability estimate of a group attack and the y-axis is the posterior estimate resulting from an adjustment by a factor of 4. Participant data show more variance and more anchoring bias (i.e.,

regression toward the posterior=prior line) than our ACT-R model. We calculated the Jensen-Shannon Divergence (JSD) of probability adjustment between the participants and our model, and the average JSD is .05. The *R* and R^2 fits are .88 and .78 respectively, which suggests our ACT-R model predicts human data closely.

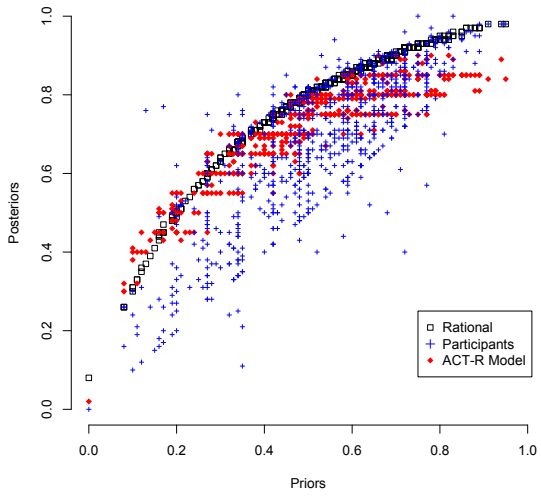

Figure 5.2 Probability adjustments for IMINT and MOVINT layers with respect to rational, participants, and ACT-R model.

6. Conclusion and Discussions

We developed an ACT-R model of sensemaking in geospatial intelligence tasks based on two widely used learning processes in ACT-R modeling: Instance-Based Learning and reinforcement learning. The model demonstrated behavior patterns similar to human participants' in terms of layer selection sequences as well as anchoring biases in probability adjustments made in response to new evidence.

6.1 Layer selection sequences

Our ACT-R model stores previous instances from Tasks 4 and 5 in its declarative chunks that capture the current situation (prior probabilities), actions (provided layers), outcome (result of the layer), and utility (distance from certainty), and develops a preference layer list based on reinforcement learning from past rewards or punishments. The preference list was used to decide on which layer should be explored first in the layer selection process, and the previous instances were used to retrieve the utility of the specific layer by blending over the prior and the layer. The retrieved utility was compared with the average utility of all layers for deciding whether to choose the layer or not.

Our ACT-R model predicts participants' layer selection sequence well, however, participants were more likely than our ACT-R model to select the layers in IMINT-MOVINT-SIGINT order, which is their vertical order or presentation in the GUI for the experiment environment. The other reason for the differences in the distribution of layer selection sequences between participants and our ACT-R model is that participants might not consider information gain at each stage of layer selection, but rather consider the troop allocation score that was given after finishing each trial after all the layer selections. That is, the SIGINT should be the first layer selection from the rational perspective because it gives the highest immediate information gain, but participants rarely chose the SIGINT layer as their first layer selection because of the potential for erroneous selection of the group probed. Finally, participants might satisfice and stick to one specific combination of layer selection more often as long as they can get a high enough troop allocation score.

6.2 Anchoring biases in probability adjustment

Our results show that when trying to retrieve an aggregated value for the adjusted probabilities based on stored chunks of past experience, the model tends to make an adjustment that is smaller (i.e., more anchored) than the rational (Bayesian) amount of adjustment. Anchoring bias seems to occur when people tend to retrieve a plausible value from their past experience instead of performing costly mental calculations. Unlike exact calculations, the retrieval will be influenced by past experiences depending on how similar they are to the current situation. Instances with prior probabilities in the mid-range are more likely to be encountered than those with extreme prior probabilities. The large amount of chunks with mid-range probabilities in declarative memory will pull the aggregated value away from the extremes in the blending process.

Our model still demonstrated less anchoring bias than the participants. This might suggest that blending is only one of the many possible mechanisms that may lead to anchoring bias. Previous studies have found that anchoring could be due to a premature satisfaction (Epley & Gilovich, 2006). That is, when people mentally adjust the value from the anchor, they stop at the end closer to the anchor, rather than the middle of the range of all plausible values. Another possible reason for anchoring in this study is that in addition to making probability adjustments, participants also need to make sure that the sum of the probabilities for the four groups is one. Thus, when this constraint is not met, participants are likely to make a second round of normalization, either mentally or with the help of the interface, without explicitly realizing that the retrieved values are already normalized (our task interface provides users the option to have the four

probabilities normalized automatically, though not all participants use that). Our previous model (Paik et al., 2012) that incorporates both blending and a second round of normalization generates produces results that are more anchored than the model presented in this paper, and had a better fit with human data. Therefore, there may be several reasons for anchoring bias in our experiment, and the ACT-R blending process is just one of these.

7. Acknowledgement

This work is supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract number D10PC20021. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained hereon are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI, or the U.S. Government.

8. References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, 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: Erlbaum.

Epley, N., & Gilovich, T. (2006). The anchoring-and-adjustment heuristic Why the adjustments are insufficient. *Psychological science, 17*(4), 311-318.

Gonzalez, C., Lerch, J. F., & Lebiere, C. (2003). Instance-based learning in dynamic decision making. *Cognitive Science, 27*, 591-635.

Klein, G. (1999). *Sources of power: How people make decisions*. Cambridge, MA: MIT press.

Klein, G., Moon, B., & Hoffman, R. R. (2006a). Making sense of sensemaking 1: Alternative perspectives. *Intelligent Systems, IEEE, 21*(4), 70-73.

Klein, G., Moon, B., & Hoffman, R. R. (2006b). Making sense of sensemaking 2: A macrocognitive model. *Intelligent Systems, IEEE, 21*(5), 88-92.

Lebiere, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft, 8*, 5-19.

Lebiere, C., Pirolli, P., Thomson, R., Paik, J., Rutledge-Taylor, M., Stazewski, J., & Anderson, J. R. (Submitted). A Functional Model of Sensemaking in a Neurocognitive Architecture

MITRE Technical Report (In Progress). IARPA's ICArUS Program: Phase 1 Challenge Problem Design and Test Specification

Paik, J., Pirolli, P., Lebiere, C., & Rutledge-Taylor, M. (2012). *Cognitive biases in a geospatial Intelligence Analysis Task: An ACT-R model.* Paper presented at the 34th annual meeting of the Cognitive Science Society, Sapporo, Japan.

Pirolli, P., & Card, S. (2005). *The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis.* Paper presented at the International Conference on Ingelligence Analysis, McLean, VA.

Reitter, D. (2010). Metacognition and multiple strategies in a cognitive model of online control. *Journal of Artificial General Intelligence, 2*(2), 20-37.

Russell, D. M., Stefik, M. J., Pirolli, P., & Card, S. K. (1993). *The cost structure of sensemaking.* Paper presented at the Proceedings of the INTERACT'93 conference on Human factors in computing systems, Amsterdam.

Rutledge-Taylor, M., Lebiere, C., Thomson, R., Stazewski, J., & Anderson, J. R. (2012). *A Comparison of Rule-Based versus Exemplar-Based Categorization Using the ACT-R Architecture.* Paper presented at the Proceedings of the 21st Conference on Behavior Representation in Modeling and Simulation, BRIMS Society: Amelia Island, FL.

Thomson, R., Lebiere, C., Rutledge-Taylor, M., Stazewski, J., & Anderson, J. R. (2012). *Understanding Sensemaking Using Functional Architectures*. Paper presented at the Proceedings of the 21st Conference on Behavior Representation in Modeling and Simulation, BRIMS Society: Amelia Island, FL.

Author Biographies

JAEHYON PAIK is a Post-Doctoral Fellow in the Augmented Social Cognition Area of the Palo Alto Research Center.

PETER PIROLLI is a Research Fellow in the Augmented Social Cognition Area of the Palo Alto Research Center.

WEI DONG is a PhD student in Information Science Department at Cornell University.

CHRISTIAN LEBIERE is a Research Faculty in the Psychology Department at Carnegie Mellon University.

ROBERT THOMSON is a Post-Doctoral Researcher in the Psychology Department at Carnegie Mellon University.