The AMBR Model Comparison Project: Round III — Modeling Category Learning

Session Organizers: **Kevin A. Gluck (kevin.gluck@williams.af.mil)**
Air Force Research Laboratory
6030 S. Kent St., Mesa, AZ 85212 USA

Richard W. Pew (pew@bbn.com)
BBN Technologies
10 Moulton St., Cambridge, MA 02138 USA

The goal of the Agent-based Modeling and Behavior Representation (AMBR) Model Comparison Project is to advance the state of the art in cognitive modeling. It is organized as a series of model comparisons, moderated by a team from BBN Technologies. In each comparison, a challenging behavioral phenomenon is chosen for study. Data are collected from humans performing the task. Cognitive models representing different modeling architectures are created, run on the task, and then compared to the collected data. The current effort focuses on models of category learning in a dynamic, dual-task environment. Model comparisons such as this, especially with directly comparable human data are rare. While models of category learning are commonplace, the fact that these are models of integrative performance, not just models of category learning in isolation, makes this set of presentations unique.

Experiment Design and Comparison of Human and Model Data

David Diller (ddiller@bbn.com)
Yvette Tenney (ytenney@bbn.com)
BBN Technologies

This experiment involved a classic concept learning task embedded in an air traffic control situation. Subjects had to learn to make correct decisions to accept or reject altitude change requests, based on three bi-variate properties of the aircraft (percent fuel remaining, aircraft size, and turbulence level). A novel feature of the experiment was the addition of multi-tasking to this concept learning paradigm. In addition to the altitude change requests (the concept learning task), the participant had to hand-off a number of aircraft to adjoining controllers (secondary task).

The design consisted of 9 conditions, defined by 3 category structures and 3 workload levels. The three category structures, borrowed from Shepard, Hovland, and Jenkins (1961), were: single attribute relevant (Type I), a single-attribute rule plus exceptions (Type III), and no rule (Type VI). The three workload levels consisted of 0, 12, or 16 required handoffs, in addition to the 16 altitude requests. It was expected that both category structure and workload level would affect performance. There were 8 scenarios, or trials, lasting ten minutes each. One hour of training on the mechanics of the tasks preceded the trials.

Ninety humans and four different human performance models described in subsequent abstracts were run through the scenarios. The interface, consisting of a radar screen with moving aircraft and action buttons, was designed to accommodate both humans and models. Humans were randomly assigned to one condition (ten per condition). The models were run one or more times in each condition.

All of the modelers were given the human learning data as soon as they were collected, and while the models were still under development. It was expected, therefore, that they would fit the data fairly well. However, a transfer test (for which the modelers were not given the human data in advance) provides an opportunity to test the generalizability of the models predictions.

Results for both humans and models will be presented on the effects of category structure and workload over trials. Human data and model data are available for the following measures: learning curves (probability of error) on the concept learning task, performance errors on the secondary task (missed and incorrect actions), reaction time on both the concept learning and secondary task, self rated workload ratings (collected from models too!), and self- reports on rule discovery and other strategies on the concept task (humans only). This presentation will set the stage for the modelers to describe the mechanisms and assumptions that allow their models to replicate the results.

An EPIC-Soar Model of Concurrent Performance on a Category Learning and a Simplified ATC Task

Ron S. Chong (rchong@gmu.edu)
George Mason University
Robert E. Wray (wray@soartech.com)
Soar Technology, Inc.

During the first phase of the AMBR project, we developed a model of a simplified en-route air traffic control task. That model was built using the EPIC-Soar architecture, an integration of the perceptual and motor systems of the EPIC architecture with Soar, a learning cognitive architecture. The task to be modeled for the current phase of AMBR is the combination of the same ATC task with a new concept acquisition task. Our approach to building the new model has been to reuse, in a modular fashion, previous Soar models for the subtasks. The ATC model is essentially the same as that of the previous AMBR phases. To produce the learning behavior, we have incorporated an existing process model of concept learning called SCA (symbolic concept acquisition). SCA was developed in Soar and has been successfully used in many Soar applications and models that require concept learning. Results of the model will be presented.

Developing Concept Learning Capabilities in the COGNET/iGEN Integrative Architecture and Associated AMBR ATC Model

Wayne Zachary (wayne_zachary@chiinc.com)
CHI Systems, Inc.

A concept learning mechanism has been added to the COGNET/iGEN modeling and human performance simulation system, and the model developed in Rounds 1 and 2 of the AMBR model competition has been extended to add the learning task. The learning mechanism developed enables learning of the conditions under which a task or goal should or should not be pursued, in addition to the current mechanism of evaluating a predefined boolean expression against contents of declarative memory. Building on the premise that concept learning can be characterized as hypothesis testing, the learning mechanism incorporates a metacognitive strategy for hypothesis generation and an hypothesis selection process based on memory for previous exemplars and their feedback, as well as previous rules or rule parameters tried, moderated by attentional and forgetting processes. The integration of the learning mechanism into the COGNET/iGEN architecture and the extended ATC model results will be presented.

An Activation-based Theory of Categorization

Christian Lebiere (cl+@cmu.edu)
Carnegie Mellon University

We propose a model of category learning implemented in the ACT-R cognitive architecture (Anderson & Lebiere, 1998). ACT-R is a hybrid architecture that combines a symbolic production system with a subsymbolic activation calculus. Our model is directly grounded in the constraints provided by the architecture, especially its declarative memory retrieval mechanism. Generalization to new instances is produced by a similarity-based partial matching mechanism that operates at the subsymbolic level. A number of latency predictions that had previously been explained in terms of a random walk process arise from an aggregate retrieval mechanism called blending. These subsymbolic mechanisms provide many of the advantages of connectionist systems while preserving inspectability at the symbolic level. The category learning model was added in a modular fashion to the existing ATC model from previous AMBR rounds.

Concept Learning: Knowing and Reasoning in the DCOG Architecture

Robert G. Eggleston (robert.eggleston@wpafb.af.mil)
Air Force Research Laboratory
Katherine L. McCreight (kate@n-spaceanalysis.com)
N-Space Analysis

DCOG is an emerging architecture of cognition. It treats cognitive behavior in terms of organized state changes that occur across a set of subsystems. Concept learning may be achieved in the architecture either by emergent knowing, derived from low-level feature-based recognition, or by higher-level reasoning using more abstract feature-derived knowledge that supports hypothesis formation and testing. In this study, a DCOG model was used to perform a complex air traffic control task that contained a concept-learning component. The concept-learning subtask was patterned after the classic Shepard, Hovland, and Jenkins (1961) task but limited to types 1, 3, and 6. Given the structure of these concept types and the balanced exemplar presentation history used in the experiment, both the knowing and reasoning pathways are viable for type 1 concepts; but only the reasoning path is viable for types 3 and 6. Both the knowing and reasoning pathways support individualistic variations or strategies and thus can emulate individual subject differences. In this presentation, we describe the feature-based concept learning infrastructure of DCOG and discuss its performance on the ATC task.

Symposium Discussant

Bradley C. Love (love@psy.utexas.edu)
University of Texas

Human category learning takes many forms, yet research in category learning has focused almost exclusively on one narrowly defined task: classification learning with no secondary task. This focus has allowed researchers to understand and explore their predictions in detail, but only within a circumscribed domain. Unfortunately, theoretical progress (as well as practical application) also demands the testing of boundary conditions. In many cases, we simply do not know how well our theories of learning generalize across task situations and induction tasks. The work presented in this symposium is an important step towards developing more general theories of learning that can make contact with human performance outside the laboratory.

Acknowledgments

AMBR Round III is sponsored by the Air Force Research Laboratory and the Office of Naval Research.

References

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

Shepard, R. N., Hovland, C. L., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs, 75*(13, Whole No. 517).