Work To Date

Project: The Cognitive Function of Theoretical Knowledge in Procedural Learning

Contract no.: N00014-89-J-1681

Principal Investigator: Stellan Ohlsson, [412] 624-7450 Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260.

Period: April, 1991 through March, 1992.

Background

The present project continuous work began under a previous project, "Knowledge-based revision of cognitive procedures in response to changing task demands: Towards a theory of the nature and function of principled knowledge". In the course of the previous project, a theory of skill acquisition was developed. The basic principles of this theory are that (a) skill acquisition consists of encoding the results of problem solving for future use, and (b) that skill acquisition consists of the correction of errors, where an error is defined as a contradiction between the outcomes of problem solving steps and prior domain knowledge. The theory was embodied in a simulation model called *the Heuristic Searcher* (HS). The HS model was used to clarify various phenomena in the learning of elementary arithmetic. This work has been published in

Ohlsson, S., Ernst, A. M., & Rees, E. (in press). The cognitive complexity of doing and learning arithmetic. *Journal of Research in Mathematics Eduation.*

Ohlsson, S., & Rees, E. (1991a). The function of conceptual understanding in the learning of arithmetic procedures. *Cognition & Instruction, 8,* 103-179.

Ohlsson, S., & Rees, E. (1991b). Adaptive search through constraint violation. *Journal of Experimental and Theoretical Artificial Intelligence, 3,* 33-42.

Ohlsson, S. (in press). Artificial instruction. A method for relating learning theory to instructional design. In P. Winne & M. Jones, (Eds.), *Foundations and frontiers in instructional computing systems*. Berlin, Germany: Springer-Verlag.

The purpoose of the present project is to apply the theory and the model to a more complex instance of skill acquisition.

Quarter 1: April through June, 1991.

For the continued work, we choose to model the acquisition of routine scientific problem solving skills. A typical example of a routine problem solving skill in organic chemistry is to derive the structural formulate, the so-called Lewis structure, for an organic molecule, given its molecular (sum) formula. This skill is basic to the study of organic chemistry and it is typically taught in the beginning of a college-level chemistry course. Analysis of several widely used textbooks revealed that a common instructional approach is to teach an overly general or incomplete version of the procedure for constructing Lewis structures through verbal instructions and then provide practice problems which require the learner to adapt the general procedure to the specifics of particular classes of molecules.

During the first quarter, we replicated this learning scenario within our simulation model. This required us to (a) translate the verbal instructions given in textbooks into code for the HS model, (b) identify the domain knowledge that students are expected to have at the time that they learn about Lewis structures, and (c) translate that knowledge into code. Finally, the resulting model had to be debugged.

Quarter 2: July through September, 1991

The chemistry model was studied in order to derive its quantitative behavioral predictions. Extensive learning experiments were run in which the model were given practice on sequences of problems taken from textbooks. Like a student, the model went through repeated trials on the practice problems until it could perform them without errors. This corresponds to simulating a single learner. The model was then reinitialized and a second series of training trials began, corresponding to the simulation of a second learner; and so on. The aggregated data from a sequence of such simulation runs correspond to the simulation of the aggregated data from an experiment with a
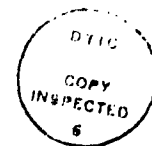
group of human learners, the most common type of learning experiment.

The results show that the model predicts the negatively accelerated learning curve that is typical of human learning. The reason for this is as follows. According to the underlying theory, skill acquisition consists of a sequence of learning events, each learning event involving the correction of one error. In an approximately uniform task environment, the amount of improvement *per learning event* is constant. However, the number of learning events *per trial* decrease as mastery is approached, because the learner makes fewer errors. Constant improvement per learning event and decreasing number of learning events per trial implies a negatively accelerated rate of improvement per trial. This is the second qualitative explanation for the negatively accelerated learning curve to be proposed[1].

Quarter 3: October through December, 1991

In order to study the learning curve in more detail, we developed a mathematical equation that describes the learning curve implied by the computational mechanisms in our model. This equation yields an exponentially decreasing curve. This result was unexpected, for two reasons. First, it has often been observed that the data from human learners follow a power law rather than an exponential law. Second, the results attained during the simulation runs in the previous quarter did, in fact, follow a power law.

In order to study this discrepancy between the simulation model and the mathematical model more carefully, we conducted a number of numerical simulations in which we aggregated exponential equations with random variation in the parameters. We found that the learning curve produced by aggregating a small number, e. g., ten, exponential learners will frequently approximate a power law quite closely. Because most learning experiments reported in the literature aggregate data across individuals in order to smooth out learning curves, this finding removes the main discrepancy between the model and existing data. However, there is also data in the literature which indicate that data sets from *single* human learners also follow

---

[1] The first qualitative explanation for the negatively accelerated learning curve was the chunking explanation proposed by Allen Newell and Paul Rosenbloom.

a power law. This discrepancy beween theory and data remains unexplained.

Quarter 4: January through March, 1992

The two major analytical techniques used in the study of learning are computer simulation and mathematical modeling. The mathematical modeling technique is limited in that it can only be applied to situations which are simple enough to be described in equations which have analytical solutions. Past work in mathematical psychology has established that this is a very restricted set of learning situations. The computer modeling technique can be applied to complex learning scenarios. However, this technique is handicapped by the amount of time and effort required to design, implement, debug, and run a complicated A. I. based model. Most such models reported in the literature have required several man years worth of work, a very labour intensive way to derive predictions from hypotheses.

In response to these problems, we began searching for a theoretical technique that would provide us with the convenience of mathematical modeling in conjunction with the fidelity to the complexity of human learning provided by computer simulation. We soon realized that the quantitative aspects of the behavior of a simulation model can be mimicked by what we call an *abstract model*. An abstract model operates upon a structure (e. g., a graph or a tree structure) that shares the quantitative properties (branching factor, path lengh, etc.) of the relevant task environment, but which does not represent the content of that environment. The abstract model mimicks the effects of learning with simple operations on that structure (e. g., removal of branches). The abstract model is a computer program and thus does not force us to find solvable equations; however, neither does it require the implementation of an AI model. The code for an abstract model is short and simple to write and runs in a fraction of the time it takes the run the AI based model.

During the fourth quarter, we perfected this technique by developing an abstract model corresponding to the HS model. We conducted two experiments on the abstract model which could not have been run on the AI based model in a reasonable time frame. In both experiments, we varied a parameter in the learning scenario and ascertained the effects on the learning curve. The two parameters were the alertness

of the learner and the amount of structure in the task environment. The results were in accord with expectations (i. e., lowered alertness yielded slower learning, increased structure in the task environment yielded faster learning), increasing our confidence in the abstract modeling technique.

Future work

Continued work will use the technique of abstract modeling to make comparative evaluations between the theory we have developed and the two alterantive theories of skill acquisition considered by researchers the present time. After completing the HS model, we will develop abstract models corresponding to the Soar model developed Allen Newell and co-workers, as well as the ACT* model developed by John Anderson and co-workers. The goal is to run the three abstract models under a variety of conditions in order to search for divergent behavioral predictions, i. e., situations in which variations in a parameter produces opposite or contrasting effects in the different models. Such situations provide a possibility to conduct empirical studies aiming to differentiate between the models. We have requested new funds to continue this line of work.