Strength Adjustment In Hierarchical Learning

Stellan Ohlsson (stellan@uic.edu)
Andrew D. Halpern (halpern@aol.com)
Department of Psychology
The University of Illinois at Chicago
1007 West Harrison Street (M/C 285)
Chicago, IL 60607, U.S.A.

Abstract

Hierarchical systems can adapt by adjusting the strengths of their components in response to environmental feedback. Regimens for propagating adjustments through a hierarchy are either *cascading* or *distributional*, depending on whether the sum of the adjustments is variable or fixed. Both types of regimens can be *dampened*, *amplified* or *sustained*, depending on whether nodes higher in the hierarchy are adjusted less, more or with the same amount as lower nodes. We show that a cascading regimen learns most efficiently with amplified propagation, while a distributional regimen learns most efficiently with sustained propagation. Cognitive scientists ought to explore a wider range of propagation regimens.

Hierarchy and Strength

Adaptive systems, whether artificial (e.g., A. I. systems), mental (e.g., skills) or natural (e.g., ecologies) tend towards hierarchical organization. In particular, cognitive skills are hierarchical in part because goals are analyzed into subgoals in the course of performing a task or solving a problem (Catrambone, 1996). For example, planning a trip might break down into deciding on mode of travel, booking accommodation and purchasing equipment, each of which breaks down into subgoals of yet smaller scope. Hierarchical task representations can also arise because a skill, once mastered, is integrated into more encompassing skills (Bruner & Bruner, 1968), a process referred to as part-whole transfer (Schmidt & Young, 1987) or vertical transfer (Gagne, 1970). For example, addition and substraction are subprocedures in the standard procedure for long division.

Adaptive hierarchical systems change in distinctive ways over time (Salthe, 1985). In particular, cognitive skills speed up during practice (Ericsson, Krampe & Tesch-Römer, 1993; Lane, 1987; Proctor & Dutta, 1995). Models of practice typically assume that components of task representations have strengths that are adjusted in response to feedback from the environment (Anderson, 1993; Logan, 1988; Ohlsson & Jewett, 1997ab). The strength values affect overt behavior primarily by resolving conflicts between competing actions.

Although the two concepts of hierarchy and feedback-driven strengthening are often invoked separately to explain learning, the cognitive literature lacks a rigorous analysis of their relation. How should feedback be propagated through hierarchical representations? The central feature of such a representation is the existence of non-terminal nodes (i.e., goals and subgoals). Goals are only indirectly linked to terminal nodes (actions) and hence to environmental input. If an action generates positive (or negative) feedback, how should the strengths of the relevant goals and subgoals be adjusted?

Cognitive models typically assume that a strength adjustment decreases in magnitude as it propagates through a network. We refer to this as *dampened* propagation. The opposite hypothesis--that the strength adjustment increases in magnitude as it propagates--is seldom studied, because such *amplified* propagation regimens are not possible in non-hierarchical networks. A signal that grows as it spreads will eventually reach all nodes in the network and increase without bounds. Amplified propagation can only work in a hierarchical representation with a well-defined stopping point (the top goal). Sandwiched between amplified and dampened strengthening is a regimen in which the strength adjustment neither decreases nor increases during propagation. We refer to this as *sustained* propagation.

Strengthening regimens also differ along another, less familiar, dimension. The propagation regimens typically considered in cognitive models are *cascading*, i.e., the change starts at some point of origin and moves to an adjacent node. The amount of change in the adjacent node is determined solely by the amount of change in the point of origin, by the relation between the two nodes and/or by some parameter. As the process moves from node to node, the change in each node is determined locally, without interaction with the changes occurring in other nodes.

Because cascading strengthening does not limit how much change can occur in the context of any one learning event, it raises questions about how the strength concept--so ubiquitous in cognitive models--should be interpreted in physiological terms. Brain researchers have not yet reached consensus on how the brain changes during practice, so we have to consider alternative hypotheses.

For example, suppose that the strength of a link between two nodes corresponds to the amount of transmittor substance available at the synapses between the relevant neural circuits. If so, then it can hardly be the case that strength increments have no upper limit. The brain is a finite physical system and it can only produce a certain amount of transmittor substance per unit of time. Hence, the amount by which a particular synapse can change in the course of a single learning event must be limited. Similar arguments apply to other physiological interpretations of the cognitive strength concept, e. g., number of synaptic connections, number of neurons, spike frequency, and so on.

Hence, it is reasonable to consider strengthening regimens in which the total amount of change in any one learning event is fixed. We call them *distributional* because they distribute the adjustments over the relevant components. Distributional regimens can be amplified (if components higher in the hierarchy receive larger adjustments), dampened (if higher components receive smaller adjustments) or sustained (if the size of the adjustment is a constant).

What are the behavioral consequences of these types of propagation regimens? How does type of propagation regimen affect learning? This is the question addressed in this paper. We present a series of computer simulations, using the abstract modeling methodology developed in past work (Ohlsson & Jewett, 1995, 1997ab). We first determine the effects of amplified, dampened and sustained propagation on the form and efficiency of learning in both cascading and distributional regimens. We then investigate the interaction between these two dimensions.

The main result is that a cascading system learns most effectively when the propagation regimen is amplified, while a distributional system learns most efficiently when the regimen is sustained, in contrast to the dampened cascade typically assumed in cognitive models.

A Hierarchical Learning System

The purpose of the computer model described in this section is to capture the essential characteristics of acquiring a hierarchical task representation via feedback during practice. The model is implemented in Macintosh Common Lisp.

Performance Module

A goal can typically be achieved by several methods; a method typically requires the satisfaction of several subgoals. Hence, a hierarchical task representation can be modeled by an AND/OR tree in which goals are modeled by OR nodes with links leading to the alternative methods and methods are modeled by AND nodes with links to the subgoals required to execute them. A task performance is a top-down, left-to-right traversal of such a tree. The terminal nodes in the tree correspond to (executable and observable) actions. The traversal of the tree generates a task solution in the form of a sequence of terminal nodes (actions). Performances can differ with respect to which method is selected at each OR node and in which order the subgoals attached to an AND node are executed and hence with respect to which sequence of actions is generated.

At each OR node, the performance module retrieves all links leading to alternative methods and selects one for execution The probability that link L is selected is a probabilistic function of its strength. Specifically, the strength of each link is multiplied with a random number between zero and unity and selects the link with the highest product. The AND-node accessed by that link is instated as the current node.

At each AND node, the performance module retrieves all links leading to conjunctive subgoals and orders these in accordance with their strengths. The probability that link L_1 is ordered before link L_2 is a probabilistic function of their strengths. Specifically, each strength is multiplied with a random number between zero and unity and the subgoals are ordered in accordance with the resulting products.

Notice that decisions are made by comparing link strengths in both OR and AND nodes. Hence, both types of decisions are affected by the propagation of strength increments.

Learning Mechanisms

We designate an arbitrarily chosen sequence of terminal nodes (actions) as the target performance. Initially, all links have strengths equal to unity. The performance module then chooses methods randomly in OR nodes and poses subgoals in random order in AND nodes, thus generating a haphazard sequence of actions. Learning how to perform the target task is to adjust the strengths such that the correct method is chosen in each OR node and the subgoals are posed in the correct order in each AND node.

During execution, each terminal node generated by traversing the AND/OR tree is compared to the corresponding action in the target performance. If they match, the model receives positive feedback. If not, the model can receive negative feedback but this feature was not used in the work reported in this paper (see Ohlsson, 1996, Ohlsson & Jewett, 1997ab, for models of negative feedback). Upon receiving feedback the model propagates strength increments upwards through the AND/OR tree.

Strength adjustment in OR nodes is straightforward: The strength of the successful link is increased with an amount determined by the particular strengthening regimen used. In AND nodes, a successful subgoal is strengthened in proportion to its position in the sequence of subgoals. A successful subgoal is strengthened more if it succeeds as the first link selected in the relevant AND node than if it succeeds as, for example, the third link selected. Successive strength increments thus orders the links (subgoals) in the AND node. In short, the model learns both to choose the appropriate method (in OR nodes) and to execute the chosen method correctly (in AND nodes).

One traversal of the AND/OR tree corresponds to one training trial. Over repeated trials, the link strengths gradually change and the model begins to make correct decisions. Eventually, the sequence of terminal elements generated matches the sequence of elements in the target performance, i.e., the model performs the target task correctly. When the model reliably generates the target performance, the task has been mastered.

Simulation Methodology

Our simulation experiments replicated the methodology of experimental studies of human skill practice: A set of learners are given a sequence of practice trials until they reach a criterion of mastery. (Due to the probabilistic nature of the decision algorithm, the behavior of the model varies from run to run.) Quantitative measures are averaged across learners.

The simulation experiments reported here used between 10 and 100 simulated learners, depending on how many learners were needed to obtain regular results. The criterion of mastery was three consecutive error-free performances of the target task. The branching factor of the hierarchical representation varied between three and six, and the height between two and eight.

The simulation results are reported in terms of two numerical measures. First, *effort per trial* measures task performance in terms of the number of decisions that the model had to make to perform the task. Each visit to a node counts as one decision. Both OR nodes and AND nodes are counted. Repeated visits to a node due to error-induced backup are counted as separate decisions. If we assume that each decision takes time, then effort per trial should be monotonically related to solution time, the measure most often used in empirical studies of practice effects.

Second, *effort to mastery* is computed by summing effort per trial across all trials needed to reach the criterion of mastery for the target task. This variable measures how efficiently the system learns.

Cascading Strengthening Regimens

Intuition suggests two opposite ways to propagate strength increments in response to positive feedback. First, because non-terminal nodes are only indirectly related to the terminal nodes (actions) that generate the feedback, it is reasonable to adjust their strengths with a smaller amount than that used to adjust the strength of the terminal node. That is, as the strength increment is propagated upwards, it becomes smaller and smaller. This is dampened propagation.

On the other hand, a wrong choice at a node high in the tree has more devastating effects on performance than a wrong choice at a low node, because all nodes dominated by the node at which the incorrect decision is made will themselves be incorrect and a higher node dominates more nodes than a lower one. This suggests that strength increments ought to increase as they are propagated upwards. This is amplified propagation.

There seems to be no *a priori* rationale for sustained propagation.

We implement all three types of regimens by multiplying the current strength (initially set to 1.0) by unity plus a strength increment (in these simulations arbitrarily set to 0.1), multiplied by a *propagation parameter* (*pp*), which itself is raised to the number of steps in the propagation process. If s_t is the strength of link X at time t and X is n levels above a terminal node that generates positive feedback, then the strength of X at time t+1 is

$$s_{t+1} = s_t * (1 + 0.1 * pp^n). \qquad (1)$$

If the parameter *pp* is smaller than unity, then the increment is smaller and smaller for each step in the propagation process, and we have dampened propagation. If *pp* is larger than unity, we have amplified propagation. Finally, if *pp* is equal to unity, we have sustained propagation

How does type of strengthening regimen affect the behavior of the model? Figure 1 shows learning curves for five values of the propagation parameter, two smaller than unity (0.3 and 0.6), unity and two larger than unity (1.2 and 1.8). All five curves are roughly linear when plotted with logarithmic coordinates, i.e., they approximate power laws, as do human learning curves (Delaney, Reder, Staszewski & Ritter, 1998; Lane, 1987; Logan, 1988; Newell & Rosenbloom, 1981). For the sustained regimen (*pp* = 1.0), shown by the straight line in Figure 1, the r^2 fit to a power law is better than 0.99. The power law fits for the dampened propagation regimens are .93 (*pp* = 0.3) and .98 (*pp* = 0.6), suggesting increasing departure from power law fit the more severe the dampening. The power law fits for amplified propagation regimens are .96 (*pp* = 1.2) and .95 (*pp* = 1.8). In short, a sustained propagation regimen generates a better fit to a power law than either an amplified or dampened strengthening regimen.

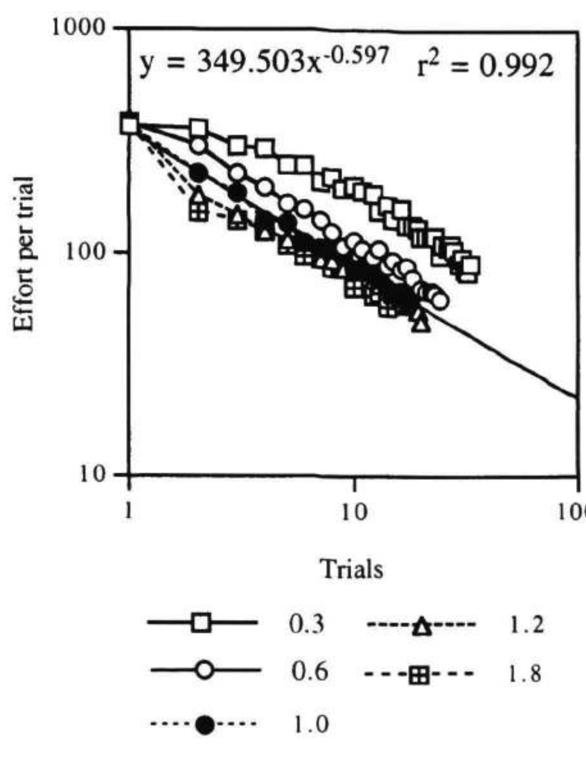

Figure 1: Learning curves for five different strengthening regimens in a cascading system, plotted with logarithmic coordinates. The straight line and the equation shows the power law fit for the sustained regimen (*pp* = 1.0).

A second observation is that the five curves in Figure 1 form an orderly progression, with gradual downward displacement for increasing values of the *pp* parameter. That is, learning appears to be become more efficient as *pp* increases.

This impression is confirmed if we plot learning efficiency in terms of total effort to mastery as a function of the propagation parameter; see Figure 2. The effort to mastery turns out to be a monotonically decreasing function of the *pp* parameter. *When strengthening cascades, amplified propagation leads to more efficient learning than dampened or sustained propagation,* although the advantage over sustained propagation quickly approaches an asymptote once $pp > 1$.

The reason for this pattern is that amplified propagation quickly settles high-level choices. The model therefore spends less time exploring the wrong part of the hierarchy.

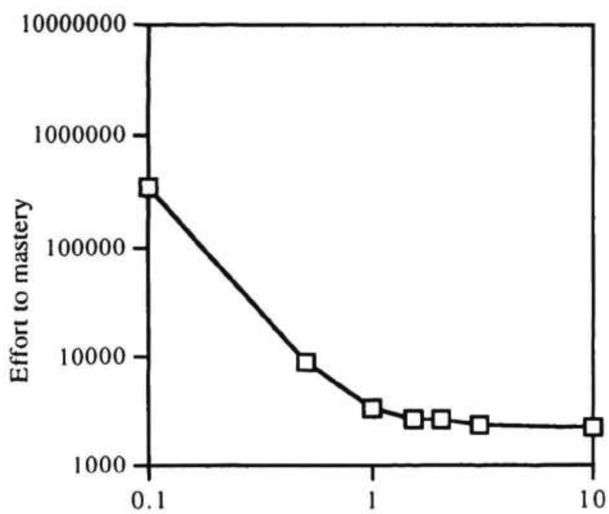

Figure 2: Learning efficiency for cascading propagation. Tree size was 3 (branching factor) by 6 (height).

Distributional Strengthening Regimens

The models described in the previous section implicitly assume that strength adjustments are *variable*, i.e., that each learning event can add a larger or smaller amount of strength to the system. In particular, the amplified regimens add more strength to the representation than the dampened regimens. The more amplified the propagation, the larger the amount of strength that is added to the representation in each learning event. This may or may not be a plausible assumption for a physically realized system like the human brain or an A. I. system.

To model learning with fixed strength adjustments, we set the amount of strength to be distributed in any given learning event to a fixed number. To make simulations comparable, we made the amount equal to the amount of strengthening distributed by the first run of the cascading system. We distributed that amount across the relevant links in such away that the relations between the strength increments conform to Equation 1 and the increments sum to the set limit.

Figure 2 shows learning curves for the distributional model. (These curves overlap too much to be displayed in a single panel.) First, panel (a) displays the curve for $pp = 1.0$. The curve has a slight S-shape across the first 5-8 trials, but the power law fit is better than .96.

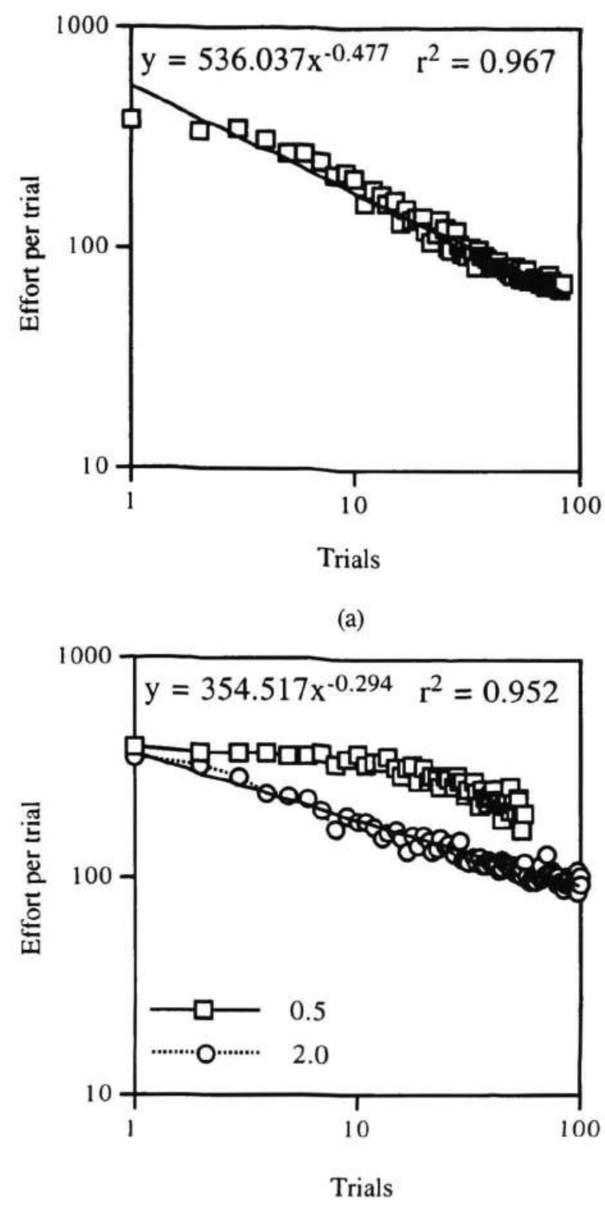

Figure 3: Learning curves for distributional regimens. Panel (a) displays the curve for a sustained regimen ($pp = 1.0$), while panel (b) displays curves for dampened ($pp = 0.5$) and amplified ($pp = 2.0$) regimens.

Panel (b) in Figure 3 shows the learning curves for a distributional regimen when propagation is dampened (*pp* = 0.5) as well as amplified (*pp* = 2.0). The amplified regimen produces a very linear curve, although the power law fit (.95) is somewhat lower than that for *pp* = 1.0 (0.99) due to the higher variability of points around the end of the curve. Thus, an amplified regiment robustly generates power law curves. The curve for *pp* = 0.5 has a slight positive curvature and its fit to a power law is the lowest of all the curves considered (0.74).

As the alert reader might have noticed, the curves in Figure 3 does not exhibit the same successive downward displacement as the curves in Figure 1, indicating that learning efficiency is not a monotonic function of the propagation parameter for distributional regimens.

This impression is confirmed by Figure 4, which displays effort to mastery as a function of the *pp* parameter. The distributional model exhibits a U-shaped pattern. Dampened propagation generate less efficient learning, but so does amplified propagation. *When strengthening is distributional, sustained propagation leads to more efficient learning than either amplified or dampened propagation.* That is, hierarchical learning is most efficient when the size of the strength increment does not change as the increment is propagated upwards through the hierarchy.

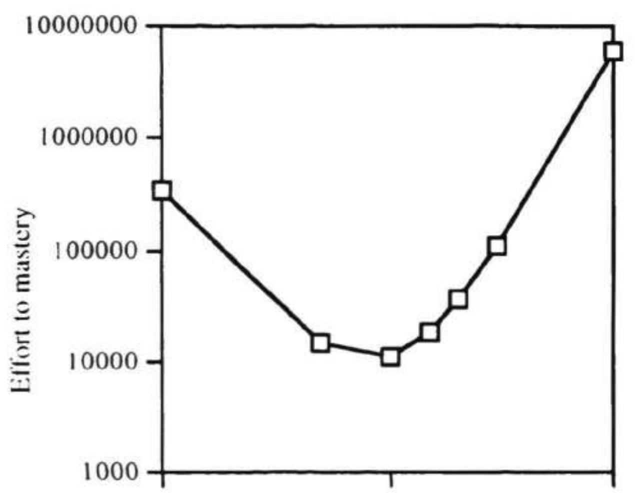

Figure 4: Learning efficiency for distributional regimens. Tree size was 3 (branching factor) by 6 (height).

The results in Figures 2 and 4 were derived for a hierarchical structure with a branching factor of three and a height of six. To bring out the contrast between cascading and distributional propagation, and to bolster our main finding, Figure 5 shows a replication for a structure with a branching factor of two and a height of eight. The pattern is replicated.

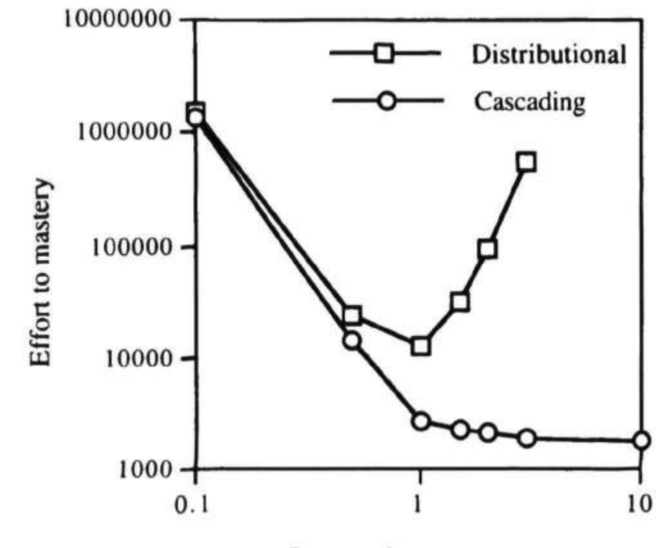

Figure 5: Efficiency of hierarchical learning for cascading and distributional strengthening as a function of propagation regimen. Tree size was 2 (branching factor) by 8 (height).

Discussion

Adaptive systems are hierarchically organized. The notion that adaptation occurs via gradual strengthening of some system components over others in response to information from the environment is ubiquitous throughout the cognitive sciences. However, it is not obvious how to combine the two ideas of hierarchy and strengthening. A hierarchical structure poses the problem of how a strength adjustment is to be propagated through the hierarchy. If a terminal node receives feedback, which of the non-terminal nodes in the hierarchy should also have their strengths adjusted and how?

In the particular class of systems we explored, cascading systems--i.e., systems that can add more or less strength to the system in each learning event depending on the propagation regimen used--produces power law learning for amplified, dampened and invariant strengthening regimens, although the power law fit lessens for severely dampened regimens. Distributional systems--i.e., systems that distribute a fixed amount of strength over the relevant links--also produces power law learning for all three types of propagation regimens. Power law fit disappears for severely dampened regimens. (In a maximally dampened regimen, i.e., a regimen in which the strength increment is not propagated at all, there is no learning and hence no learning curve.) In short, the shape of the learning curve is largely unaffected by the strengthening mechanism.

In contrast, the efficiency of learning, i.e., the effort required to reach mastery, is strongly affected. In a cascading system, learning efficiency is greater for amplified than for sustained and dampened propagation. In a distributional system, on the other hand, learning is maximally efficient for sustained propagation.

The results confirm one conclusion reached in prior work on non-hierarchical learning: That, contrary to claims sometimes made on its behalf (Logan 1988), the power law of learning has little power to discriminate between learning theories (Ohlsson & Jewett, 1997ab). (A similar conclusion has been reached by Shrager, Hobb & Huberman, 1988, on different grounds.)

The efficiency results were unexpected and more interesting. They become more interesting still if we assume that humans have evolved to be optimal learners. Natural selection does not necessarily generate optimal designs, but it does so sometimes and our capacity for learning is the key stone of our species' survival strategy. Hence, it is at least possible that we are optimal learners.

In conjunction with this assumption, our first set of results (Figure 2) imply that if the human brain operates with a cascading regimen, we should find that strength adjustments are amplified as they are propagated. On the other hand, if the brain operates with a distributional regimen, we ought to find that strength adjustments are propagated without increasing or decreasing in magnitude (Figure 4). These implications are interesting because cognitive models are typically using dampened propagation.

To discover and investigate quantitative regularities like the ones we report on this paper, we need models that allow us to make *comparative evaluations* (e.g., between cascading and distributional models) and *sensitivity experiments* (Schneider, 1988) in which we systematically vary parameters (e.g., the propagation parameter). We have found abstract computer models to be a better tool for this kind of exploration than symbolic models (Ohlsson & Jewett, 1995, 1997ab), as have others (Cooper, Farrington & Shallice, 1996).

It would be premature to make strong claims for our conclusions. The present results need to be replicated and generalized by varying the decision algorithm, the density of feedback, the tree representation and other variables. The only claim that we can make at this point is that amplified and sustained propagation within distributional strengthening regimens deserve to be considered in models of hierarchical adaptive systems. Whether this message will itself be amplified or dampened as it propagates through the cognitive science community remains to be seen.

Acknowledgments

Preparation of this paper was supported by Grant No. N00014-97-1-0826 from the Cognitive Science Program of the Office of Naval Research (ONR).

References

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Bruner, J. S., & Bruner, B. M. (1968). On voluntary action and its hierarchical structure. *International Journal of Psychology*, *3*, 239-255.

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 1020-1031.

Cooper, R., Fox, J., Farringdon, J., & Shallice, T. (1996). A systematic methodology for cognitive modelling. *Artificial Intelligence*, *85*, 3-44.

Delaney, P. F., Reder, L. M., Stasewski, J. J., & Ritter, F. E. (1998). The stategy-specific nature of improvement: The power law applies by strategy within task. *Psychological Science*, *9*, 1-7.

Ericsson, K. A., Krampe, R.Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, *100*, 363-406.

Gagne, R. M. (1970). *Conditions of learning* (2nd ed.). London, UK: Holt, Rinehart & Winston.

Lane, N. (1987). *Skill acquisition rates and patterns*. New York, NY: Springer-Verlag.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492-527.

Newell, A., & Rosenbloom, P. (1981). Mechanisms of skill acquisition and the law of practice. In J. Anderson (Ed.), *Cognitive skills and their acquisition* (pp. 1-55). Hillsdale, NJ: Erlbaum.

Ohlsson, S. (1996). Learning from performance errors. *Psychological Review*, *103*, 241-262.

Ohlsson, S., & Jewett, J. J. (1995). Abstract computer models: Towards a new method for theorizing about adaptive agents. In N. Lavrac & S. Wrobel (Eds.), *Machine Learning* (pp. 33-52). Berlin, Germany: Springer Verlag.

Ohlsson, S., & Jewett, J. (1997a). Simulation models and the power law of learning. In M. G. Shafto & P. Langley (Eds.), *Proceedings of the 19th Annual Conference of the Cognitive Science Society* (pp. 584-589). Mahwah, NJ: Erlbaum.

Ohlsson, S., & Jewett, J. J. (1997b). Ideal adaptive agents and the learning curve. In J. Brzezinski, B. Krause & T. Maruszewski (Eds.), *Modelling in psychology* (pp. 139-176). Amsterdam, The Netherlands: Rodopi.

Proctor, R. W., & Dutta, A. (1995). *Skill acquisition and human performance*. Thousand Oaks, CA: SAGE.

Salthe, S. N. (1985). *Evolving hierarchical systems*. New York: Columbia University Press.

Schmidt, R. A. & Young, D. E. (1987). Transfer of movement control in motor skill learning. In S. M. Cormier & J. D. Hagman (Eds.), *Transfer of learning* (pp. 47-79). New York: Academic Press.

Schneider, W. (1988). Sensitivity analysis in connectionist modeling. *Behavior Research Methods, Instruments, & Computers*, *20*, 282-288.

Shrager, J., Hogg, T., & Huberman, B. A. (1988). A graph-dynamic model of the power law of practice and the problem-solving fan-effect. *Science*, *242*, 414-416.