

THE "ARTIE" SIMULATION OF OPERANT CONDITIONING

LA SIMULACIÓN "ARTIE" DEL CONDICIONAMIENTO OPERANTE

MATT J. MORRIS¹
EMORY UNIVERSITY

ABSTRACT

Computational modeling is emerging as a new method for researchers to test theories of behavior. Computational models that incorporate adaptive networks are especially attractive to behavior analysts because of the selectionist nature of these models. A novel adaptive network model of operant conditioning, called "Artie," is presented. Results from simulated laboratory experiments with Artie replicated laboratory results from real animals, lending support to selectionist accounts of operant conditioning. In general, adaptive network models of behavior are probably better suited for testing theories of behavior than for teaching us about the functioning of the brain.

Keywords: Artie, behavior analysis, computational modeling, behavior theories, adaptive-network models, selectionist accounts

RESUMEN

El modelamiento computacional está emergiendo como un método nuevo para que los investigadores pongan a prueba teorías de la conducta. Los modelos computacionales que incorporan redes adaptativas son especialmente atractivos para los analistas conductuales, debido a la naturaleza seleccionista de estos modelos. Se presenta un nuevo modelo de redes adaptativas del condicionamiento operante llamado "Artie". Resultados de experimentos de laboratorio simulados con Artie replicaron resultados obtenidos con animales reales, dando así apoyo a las explicaciones seleccionistas del condicionamiento operante. En general, los modelos de redes adaptativas de la conducta son probablemente mejores para poner a prueba teorías de la conducta que para enseñarnos acerca del funcionamiento del cerebro.

¹ VerbalTech Labs, 2955 Valmont Rd., Suite 300. Boulder, CO 80301, USA. E-mail: mattjmorris@earthlink.net

Palabras clave: Artie, análisis conductual, modelamiento computacional, teorías de la conducta, modelos de redes adaptativas, aproximaciones seleccionistas

Behavior analysts have begun to recognize the power of computational models as well as the compatibility between certain types of computational models and behavior analysis (Hutchison, 1984, 1997). Some behavior analysts are becoming especially interested in computational models that utilize adaptive networks (e.g., Donahoe & Palmer, 1989; Kehoe, 1989; Stephens & Hutchison, 1993). Adaptive networks are selectionist models, meaning that their operations can be characterized by a repeating cycle of variation, selection, and retention (Palmer & Donahoe, 1992; Donahoe & Palmer, 1994). Operant learning has also been characterized as a selectionist process (Skinner, 1981; Smith, 1983, 1986), theoretically making it possible to simulate the process of operant conditioning using an adaptive network.

The current paper provides a brief introduction to the computational modeling approach to the study of animal behavior. Included in the introduction are discussions of the general simulation of animal behavior, the simulation of operant behavior using adaptive networks, and the concept of "selectionism." Afterwards, a novel computational model of operant behavior is presented. The model includes a computer-simulated organism named "Artie," whose behavior is controlled by an adaptive network. Attempts to produce free-operant learning phenomena with Artie provided tests of the model's ability to simulate operant conditioning, and the results of the simulations are presented. Lastly, the use of results from this and other adaptive network models is discussed.

The simulation of animal behavior

Computer models of animal behavior can be classified into two main categories: models that simulate group behavior and models that simulate the behavior of individual organisms. To demonstrate the utility of computer models of animal behavior, four models are presented. Two of the models concentrated on group behavior, and two of the models focused on individual behavior. Each study demonstrates the ability of computer simulations to help researchers explore real-world animal behavior.

Group behavior

Reynolds (1987) created a realistic simulation of the flocking behavior of groups of birds by programming a few basic rules thought to underlie the flying behavior of individual birds. The rules were applied only to individual

birds, with no rules applied to the group as a whole. When the behavior of the individual birds interacted, flocking emerged. By examining the parameters of the individual rules and how they interacted, Reynolds and other researchers have been able to better understand how flocking may occur in nature (Rucker, 1993, pp. 67-75).

Langton (1986) created one of the earliest simulations of group behavior. Utilizing cellular automata (an array of cells arranged in checkerboard fashion, with each cell's behavior being determined by the behavior of its neighbors), Langton created a population of virtual ants, or "vants." Only simple rules were used to simulate the behavior of each individual vant; however, when the vants interacted through the use of simulated pheromones, a self-organization very similar to that found in colonies of social insects emerged. Whether or not the rules Langton used are indeed the same rules that real insects employ, simulations such as these help researchers to understand how complexity in the natural environment can emerge from simple processes (Levy, 1992, pp. 93-107).

Individual behavior

The second computer simulation approach to studying animal behavior is the simulation of the behavior of individual organisms. Studies of individual behavior can be classified into those that require experimenter intervention and those that allow the organism to exist autonomously in an environment with which it interacts. Kemp and Eckerman (1995) have argued that autonomous organism simulations, or "in-situ" simulations, as they term them, allow for stricter testing of theories of behavior. Reviewed below are two noteworthy studies of this type.

Beer (1990) was one of the first researchers to study animal behavior using the autonomous organism approach. Beer simulated the complex behavior characteristic of several species of insects with a series of "computer cockroaches." Aided by biologically-inspired adaptive networks, his insects were able to learn to use their six legs to walk, and even to forage for food in a two-dimensional environment. Then, by destroying various connections in the neural networks, Beer was able to replicate the effects of lesion studies performed on real insects.

Terzopoulos, Tu, and Grzeszczyk (1994) also applied the autonomous-organism approach to studying animal behavior. These researchers simulated individual fish living in a virtual three-dimensional environment. The fish were developed with a realistic set of piscine muscles, fins, low-level optical sensors (eyes), and a neural network. The fish were able to learn from "scratch" how to "swim" and how to navigate their complex environment, adhering to the

biomechanics of real fish and the laws of hydrodynamic locomotion. They even learned to interact with other virtual fish and to perform behaviors such as mating, foraging, and predator avoidance. By simulating these behaviors, the authors were able to explore some of the rules that may allow real fish to produce these behaviors.

Behavior analysts have also created simulations of the behavior of individual organisms. A special class of these simulations model operant conditioning with adaptive networks. Before this class of simulations is reviewed, it should be mentioned that notable behavior analytic computer models exist that simulate operant conditioning without using adaptive networks (e.g., Shimp, 1969; Silberberg & Zirix, 1985). Likewise, some behavior analytic simulations use adaptive networks but do not simulate operant conditioning (e.g., Burgos, 1997; Moore & Choi, 1997). Only adaptive network studies of operant conditioning will be considered in the current paper.

Adaptive network models of operant conditioning

Stephens and Hutchison (1992) created a sophisticated autonomous organism known as "7G," whose behavior was determined by an adaptive network. 7G learned to display several basic operant phenomena, such as acquisition, extinction, conditioned reinforcement, and stimulus control. Even more remarkably, 7G was able to learn a variety of verbal behavior skills, such as imitation, naming, rule learning, and rule following. Because the authors employed strictly behavior analytic principles in the training of 7G, they were able to provide a sufficiency proof for a contingency approach to verbal learning, concurrently disproving one of Chomsky's (1959) claims about the insufficiency of such an approach.

Donahoe, Burgos, and Palmer (1993) created a novel adaptive network that also learned from experience based on behavior analytic principles. The experimenters were able to train the network to show such learning phenomena as acquisition, extinction, reacquisition, conditioned reinforcement, and stimulus control. Distinguishing it from most other behavior analytic adaptive networks, the adaptive network in this study was developed based upon known principles of neurobiology. The authors point out that maintaining "biological plausibility" in behavioral adaptive network models can help integrate behavior analysis with neuroscience. However, other researchers have disputed the necessity or even advantage of using biologically faithful adaptive networks when studying phenomena that occur at the behavioral level (e.g., Hutchison, 1997; Marr, 1997).

More recently, Spier and McFarland (1998) created a simulated organism existing in a virtual Skinner box with an adaptive network determining its

behavior. The simulation was used to test the theory that an animal must be employing internal representations when its preference for one of two concurrently available reinforcers is decreased outside of the experimental apparatus and it shows an immediate decrease in responding for just the "devalued" reinforcer when placed back in the experimental apparatus. The authors created their organism without the ability to form internal representations. Nevertheless, when the organism was run on the simulated procedure, it immediately decreased its responding for the devalued reinforcer but not its responding for the non-manipulated reinforcer. Once again a sufficiency proof was produced by a simulation study, this time questioning how "cognitive" animals must be in order to exhibit certain behavioral phenomena.

The above three studies give some indication of the use of adaptive network studies. One of the main reasons that behavior analysts are interested in adaptive networks is that adaptive networks are selectionist models. "Selectionist" is a concept derived from selection by consequences theory, which is reviewed here. Palmer (1997) and Palmer and Donahoe (1992) provide more extensive reviews of the relation of selectionism to adaptive networks.

Selection by Consequences

"Selection by consequences" was the name Skinner gave to the theory that a parallel exists between operant conditioning and natural selection (Skinner, 1986, 1981). More specifically, selection-by-consequences theory posits that operant conditioning is responsible for the ontogeny of behavior during the lifetimes of individual organisms in a manner analogous to the manner in which natural selection is responsible for the evolution of species over generations (Donahoe, Burgos, & Palmer, 1993; Smith, 1983, 1986). To evaluate this claim, operant conditioning and natural selection must be understood within a common framework. This may be done by describing both as three-step iterative processes.

The Process of Natural Selection

Campbell (1960), and others (e.g. Plotkin, 1994) have described natural selection as a process with three repeating steps.

- 1) variation of traits within a population
- 2) selection of certain traits over others based on the differential survival and reproduction conferred by those traits
- 3) retention of traits across generations (by the mechanism of heredity)

Whenever the three steps of natural selection operate on a distribution of traits that has not already achieved equilibrium, there is an increase in frequency of certain traits relative to others in the same population. (Ridley, 1993). A real-world example of the three steps of natural selection causing trait frequency shifts in a population comes from the work of Peter and Rosemary Grant (Weiner, 1994). The Grants studied species of Finches in the Galapagos Islands, and found that the beaks of these birds varied in size and that beak size was hereditary. When a change in the ecology of these birds caused small beaks to be favored over large ones, the distribution of beak sizes shifted towards the lower end; a larger percentage of Finches were born with smaller beaks than in previous years (Weiner, 1994, pp.100-104). Endler (1986) summarizes other naturalistic studies that have shown changes in frequencies of traits as a result of the three steps of natural selection.

The three steps of natural selection occur as an iterative process, repeating themselves with each new generation. Evolution is the cumulative change in the traits of organisms over many generations. The cumulating of changes over generations is what allows natural selection to create such wonderfully complex conglomerates of traits (usually referred to as "organisms"). If the theory of selection by consequences is correct, then the cumulating of changes is also what allows operant conditioning to create complex behavioral repertoires (Glenn & Madden, 1995).

The Process of Operant Conditioning

Like natural selection, operant conditioning may be described as a three step process:

- 1) variation of operants within an operant class
- 2) selection of certain operants over others based on the differential reinforcement of those operants
- 3) retention of operants across time (by means of the central nervous system)

Whenever the above three steps occur, certain operants become more frequent relative to others in the same operant class. An example of changes in operant distributions resulting from reinforcement come from experiments in which a rat or a pigeon in a Skinner box was required to press a lever or a key with some minimum force in order to obtain reinforcement (e.g., Stokes, 1995; Cole, 1965; Notterman, 1959). In experiments such as these, there is always variability in the force applied by the organism, and when the environment is constant, the organism's responses are fairly consistent. But when the amount of force that is required to obtain reinforcement is increased or decreased, the distribution of forces applied by the organism shifts in the direction of the new

requirement. For example, when the force requirement is increased, the organism's responses, on average, become more forceful.

Like natural selection, operant conditioning is an iterative process, allowing changes in responses to cumulate over time. As long as there is an operant contingency in effect, the three steps of operant conditioning repeat themselves, building on the effects of previous iterations. As a result, complex behavioral patterns "evolve" during the lifetimes of organisms.

METHOD

Simulation Specifics

A computer-simulated organism named "Artie" existed in a two-dimensional simulated environment. Artie and his environment were displayed graphically on the monitor screen of computers running simulations. The graphic representation of Artie contained a head, torso, two arms, and two legs. The environment represented an operant chamber, or "Skinner box," containing four walls and either one or two operant levers. The dimensions of the Skinner box were 133 pixels long by 95 pixels wide. The dimensions of each operant lever were 20 pixels long by 1 pixel wide.

The first operant lever ("operant lever #1") was placed in the Skinner box vertically with its top left corner 25 pixels from the left wall and 70 pixels from the top wall. When the second operant lever ("operant lever #2") was placed in the Skinner box for Experiment 3, it was placed vertically with its top left corner 22 pixels from the right wall and 70 pixels from the top wall. In order for Artie to "depress" operant lever #1, his right hand had to pass through the graphic representation of the lever, from right to left. In order for Artie to "depress" operant lever #2, his left hand had to pass through the graphic representation of the lever, from left to right. At the beginning of every experiment, Artie was placed with his torso seven pixels to the left of operant lever #1, making it likely that he would press that lever early in the experiment even if behaving randomly.

Although Artie could occupy any pixel position in the Skinner box, the floor of the box was divided into 30 equally spaced cells for the purpose of giving Artie feedback regarding his position. In addition to moving around his environment, Artie could place each of his arms in one of seven positions, with each position causing the respective hand to be a different distance away from his torso. Artie's arm span varied from 10 pixels when fully contracted to 38 pixels when fully extended.

Apparatus

The Microsoft Visual Basic 3.0 programming language was used to create the simulation and all associated data analysis programs. All computer programs were later upgraded to Microsoft Visual Basic 5.0. The simulation was run on standard Pentium 133 MHz PC computers.

Procedure

All simulations with Artie were run as a series of iterations, or "time steps." Each time step, the adaptive network was presented with a set of inputs indicating the current state of Artie's environment. Based on these inputs, the adaptive network produced a set of outputs that were used to select actions. Artie performed the actions, which changed the state of the environment and produced feedback in the form of reinforcement or punishment. Afterwards, a new time step began and the process repeated itself.

The adaptive network controlled all relations between Artie's environment and his behavior. To explain the operation of the adaptive network, first the inputs and outputs of the network are described. Second, the structure of the network and the method for selecting outputs based on the inputs is explained. Finally, the manner in which feedback from the environment altered input-output relations in the network is discussed.

Inputs and outputs

Every time step, three scalar values representing the current state of Artie's environment were delivered to the network. These "input values" represented Artie's current floor position in the Skinner box, the position of his left arm, and the position of his right arm. In addition, five output values determining Artie's actions for the current time step were output by the network every time step. These "output values" were used to select Artie's movement in the vertical plane of his environment, his movement in the horizontal plane of his environment, the positioning of his left arm, the positioning of his right arm, and the amount of variation that was added to the network. This last output value affected the amount of variation that would be seen in Artie's behavior. Note that while variation may not fit with the traditional notion of an "action," it is consistent with the findings of Page and Neuringer (1985) that variation in behavior can itself be conceptualized as a behavior.

Network structure and selection of outputs

The adaptive network translated inputs into outputs via two sub-networks. Both of the sub-networks were fully connected two-layer adaptive networks. The first sub-network, referred to as the "multiplier network," used the three inputs from the environment to determine "multiplier values" that were then utilized in the second sub-network. Through the multiplier values, the multiplier network influenced the effect each type of environmental input would have on the selection of Artie's actions. The second sub-network, referred to as the "action network," used the current state of the environment, along with the multiplier values from the first network, to determine Artie's actions. Explained first is the operation of the multiplier network, and explained second is the operation of the action network.

The multiplier network contained a set of input units for each of the three types of environmental inputs. Only one unit could be active within any input set at a time, with active units assigned a value of 1 and inactive units assigned a value of 0. The multiplier network contained a set of output units for each of the five multipliers that would be utilized in the action network – one for each of the five types of environmental inputs. There was no limit on the number of output units that could be active at one time, nor on the values these units could assume.

The values of the output units of the multiplier network were determined by the equation:

$$o_j = \sum_{i=1}^n a_{ij} w_{ij} \mu_i \quad (1)$$

where o_j denotes the value of output unit j , a_{ij} represents the value of input unit i connected to unit j , w_{ij} denotes the "connection weight" between input unit i and output unit j , and μ_i represents the multiplier value utilized for the input set associated with input unit i during the previous time step. Once the values of the output units were determined, they were used stochastically to select a multiplier value from the set {0.0, 0.1, 0.3, 0.5, 0.7, 0.9, 1.0}. One value was selected from the complete set of values for each of the three input sets, with the probability of choosing a value proportional to the relative magnitude of the corresponding output value.

Like the multiplier network, the action network contained a discrete set of input units for each type of environmental input, with the ability for only one unit at a time to be active within each input set. And similar to the multiplier network, active input units were assigned a value of 1 and inactive inputs units were assigned a value of 0. The action network had a discrete set of output units for each of the five types of actions Artie performed every time step. Artie had seven alternatives to choose from for each type of action. There

were seven choices for how much he could move in the vertical plane, seven choices for how much he could move in the horizontal plane, seven positions he could place his left arm in, seven positions he could place his right arm in, and seven different variation values. Likewise, each set of output units in the network contained seven individual units – one for each response alternative. There was no restriction on the number of output units that could be active at one time, nor was there a restriction on the values these units could assume.

The values of the output units of the action network were determined by the equation:

$$o_j = \sum_{i=1}^n a_{ij} w_{ij} \mu_i \quad (2)$$

where o_j denotes the value of output unit j , a_{ij} represents the value of input unit i connected to unit j , w_{ij} denotes the “connection weight” between input unit i and output unit j , and μ_i represents the multiplier value determined by the multiplier network, during the current time step, for the input set to which input unit i belonged. Once the values of the output units were determined, they were used stochastically to select actions that Artie performed during the current time step. One action was chosen from among the seven alternatives within each action type, with the probability of choosing an action proportional to the relative magnitude of the corresponding output unit value.

Update of weights

Although developed independently, the weight update method used in the Artie simulation is very similar to Rich Sutton and Andrew Barto’s “one-step Sarsa” version of their “temporal differences” method for adapting control systems (such as adaptive networks) based on environmental feedback (Sutton & Barto, 1998). In the Artie simulation, environmental feedback occurred in the form of reinforcement, which was always a positive value, and punishment, which was always a negative value. The reinforcement and punishment values were added together to form a single “reward value,” R .

After calculating the reward value, the connection weights associated with selected outputs from the current time step in both the multiplier network and the action network were updated. Restricting the updating to only those connection weights associated with selected outputs ensured that only those weights that were responsible for the current actions received “credit” for the results of the actions. The equation for updating the connection weights was

$$w_{ij} \leftarrow w_{ij} + V \frac{|w_{ij}|}{\sum_{k=1}^n |w_{ik}|} (R + P) \quad (3)$$

where the arrow means that the value on the left is replaced with the value on the right, w_{ij} represents the connection weight from input unit i to "selected" output unit j of the particular network, V denotes the variation value selected during the current time step, R denotes the reward value, and P is the "Premack" value. The proportion value within the first set of brackets in Equation 3 distributed the reward value among the weights, based on the relative magnitude of each weight's value.

The Premack value, so named because of its relation to the "Premack Principle," (Premack, 1961), converted the difference in value between successive actions into a single reward-like value. The Premack value was calculated exclusively from values within the action network, based on the values of output units that were associated with selected actions. The equation used to calculate the Premack value was

$$P = \frac{\sum_{j=1}^n o_j(t) - \sum_{j=1}^n o_j(t-1)}{R} \quad (4)$$

where $o_j(t)$ represents the value of the j^{th} "selected" output unit in the action network during the current time step, and $o_j(t-1)$ represents the value of the j^{th} "selected" output unit in the action network during the previous time step. Note that if the summed "selected" output unit values were larger during the current time step than during the previous time step, the Premack value was positive. If the opposite was true, the Premack value was negative. The one restriction applied to the P value was that its absolute value was not allowed to be larger than the reinforcement value in effect for the current experiment.

After the weights were updated by Equation 3, two operations were performed to ensure that the connection weights did not become unmanageably large or small. The first of these operations was performed every time step after the connection weights were updated and involved subjecting all weights to the transform

$$w_{ij} = \begin{cases} -10,000 & \text{if } w_{ij} < -10,000 \\ w_{ij} & \text{if } -10,000 \leq w_{ij} \leq 10,000 \\ 10,000 & \text{if } w_{ij} > 10,000 \end{cases} \quad (5)$$

The second operation was performed every 10,000 time steps, at which time all connection weights were divided in half. Although it was never explicitly tested, it was theorized that in addition to helping control the size of the weights, this procedure would simulate the effects of food depriving a real animal, making the effects of that primary reinforcer more powerful.

Now that the Artie simulation has been explained, the simulated experiments conducted with Artie and their results are reported here. Each of the experiments was designed to test Artie's ability to produce real-animal

operant phenomena, and results are presented in formats that allow their comparison to results from real animals.

EXPERIMENTS AND RESULTS

Experiment 1: Acquisition, Shaping, Extinction, and Reacquisition

Experiment 1 consisted of four phases that tested the effect of reinforcement and punishment contingencies on Artie's lever pressing behavior. Operant lever #1 was placed in the Skinner box at the previously stated position. The availability of reinforcement for depressing the operant lever varied across phases. A punishment value, on the other hand, was always delivered for lever pressing. This punishment value simulated the aversive component that has been theorized to accompany all operant responses (McDowell, Bass, & Kessel, 1993). Lever pressing was recorded in the form of cumulative records, with number of responses on the ordinate and passage of time steps on the abscissa.

Phase 1 tested Artie's ability to acquire lever pressing in the absence of explicit shaping for lever pressing. The connection weights in Artie's adaptive network were initialized to 200, a practice utilized throughout the experiments whenever the effects of previous learning needed to be erased and/or it was desired that Artie begin a procedure behaving randomly. Phase 1 lasted for 100,000 time steps, during which a reinforcement value of 200 and a punishment value of -10 were delivered for every lever press Artie produced.

Figure 1 displays the cumulative record of Artie's lever pressing during acquisition. During the 100,000 time steps of this phase, Artie pressed the lever a total of 2,400 times. Initially, lever pressing was exhibited at a low rate, but the rate soon increased as evidenced by the positively accelerated response curve. During the second half of the phase, at which point Artie seemed to have achieved a steady state of responding, he pressed the lever a total of 1,887 times for an average rate of about 38.0 lever presses per 1,000 time steps.

Phase 2 tested the effects of explicit shaping on Artie's lever pressing behavior. All of the connection weights in the adaptive network were reset to 200, erasing the effects of Phase 1. For 10,000 time steps, a reinforcement value of 50 was delivered every time Artie moved closer to the operant lever, and a punishment value of -25 was delivered every time Artie moved further away from the operant lever. Afterwards, for 100,000 time steps, a reinforcement value of 200 and a punishment value of -10 were delivered for every lever press Artie produced.

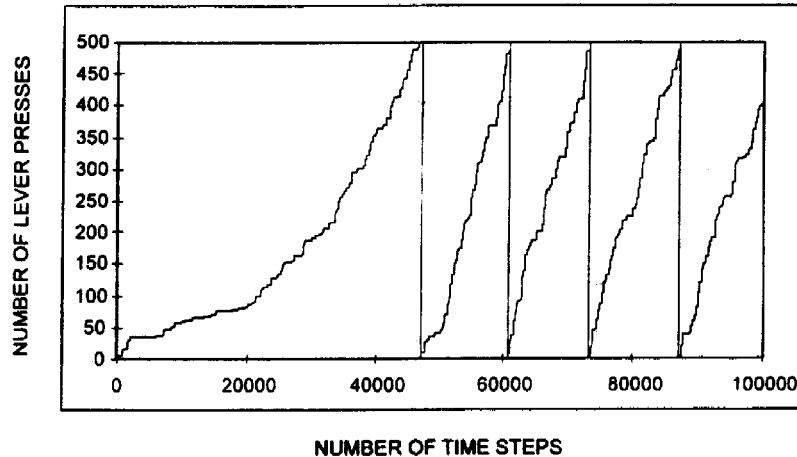


Figure 1. Cumulative record of acquisition in response to continuous reinforcement for lever pressing

Figure 2 displays the cumulative record of Artie's lever pressing after shaping. This cumulative record shows a very high rate of responding that began to slow down towards the end of the phase. In 100,000 time steps, Artie pressed the lever 10,962 times, equaling an average rate of about 109.6 lever presses per 1,000 time steps. During the last 20,000 time steps, Artie's lever pressing had slowed to an average rate of about 77.2 lever presses per 1,000 time steps. Based on Artie's high rate of lever pressing, it was decided that the connection weight values existing in the adaptive network at the end of this phase would be saved in a separate file and would be used as the starting connection weights for future experiments.

Phase 3 tested the effects of extinction on Artie's lever pressing. The connection weights in the adaptive network were not reset, in order to maintain the effects of learning during Phase 2. For 100,000 time steps, a reinforcement value of 0 and a punishment value of -10 were delivered for every lever press.

Figure 3 displays the cumulative record of Artie's lever pressing during extinction. Initially, the rate of responding was very high, at about 142 lever presses per 1,000 time steps for the first 7,000 time steps. After the initial high rate, the cumulative record shows a negatively accelerated response curve, followed by a complete cessation in responding. During the last 40,000 time steps not a single lever press was produced, equaling a rate of 0.0 lever presses per 1,000 time steps.

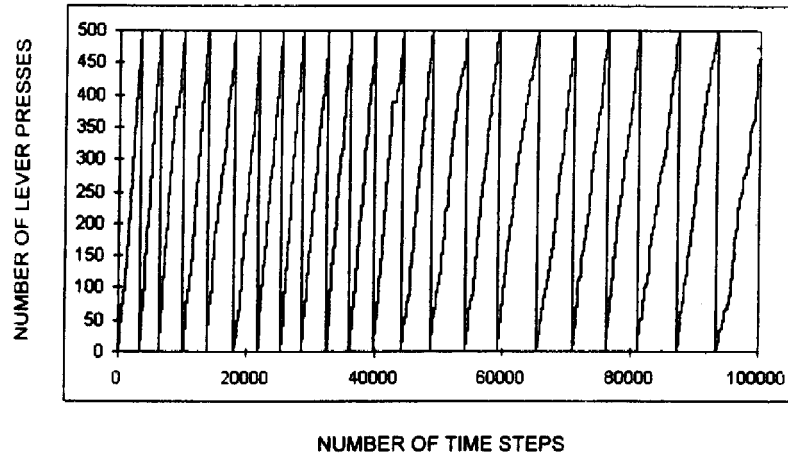


Figure 2. Cumulative record of acquisition of lever pressing after Artie was shaped to stand close to the lever

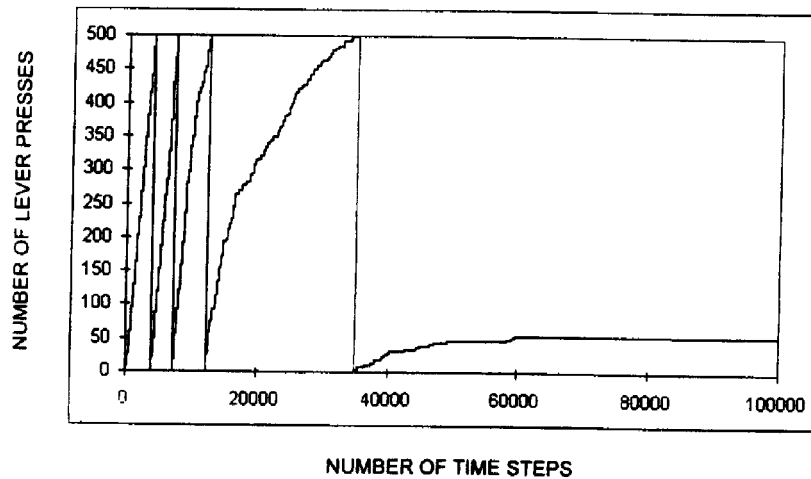


Figure 3. Cumulative record of extinction of lever pressing

Phase 4 tested Artie's ability to reacquire lever pressing after extinction. The connection weights in Artie's adaptive network were not changed, allowing the effects of learning during extinction to remain. Reinforcement was reinstated for lever pressing. For 100,000 time steps a reinforcement value of

200 and a punishment value of -10 were delivered for every lever press.

The cumulative record of Artie's lever pressing during reacquisition is displayed in Figure 4. Artie relearned to press the lever, achieving 3,136 lever presses during the 100,000 time steps. The overall structure of this cumulative record appears similar to the initial acquisition cumulative record, displayed in Figure 1, with the exception that the initial positively accelerated response curve is steeper in the current figure.

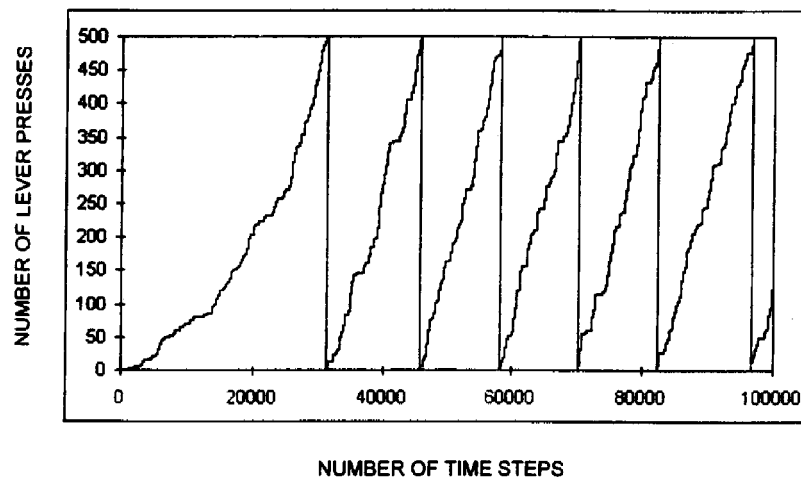


Figure 4. Reacquisition in response to continuous reinforcement for lever pressing

Experiment 2: Variable-Ratio and Variable-Interval Series

Experiment 2 tested the effects of variable-ratio (VR) and variable-interval (VI) schedules of reinforcement on Artie's lever pressing. Operant lever #1 was placed in the Skinner box and Artie was first run on a VR series and then on a VI series. Before each new VR or VI schedule, the connection weights in Artie's adaptive network were reset to the values that were saved after Phase 2 of Experiment 1.

Phase 1 tested the effects of a series of VR schedules on the rate of Artie's lever pressing. Artie was run on schedules VR 5, VR 10, VR 20, VR 30, VR 40, and VR 50, with each schedule lasting 300,000 time steps. The reinforcement value was set to 600 and the punishment value to -10 for each schedule. The average rates of responding from the last 10,000 time steps of each schedule were recorded and then graphed.

The top panel of Figure 5 displays the graph of Artie's parametric VR lever pressing behavior. In order to compare the graph displaying Artie's behavior to those from animal experiments, it was necessary to translate between time steps and seconds. It was decided that one time step would be considered roughly one second. The ordinate of the graph displays the average rate of responding per minute (60 time steps) and the abscissa displays the average number of obtained reinforcers per hour (3,600 time steps). Note that the data is fit well by a hyperbolic function, with approximately 92.7% of the variance accounted for. The structure and fit of this data is similar to that found with real animal data (see McDowell & Wixted, 1988) for data from real organisms).

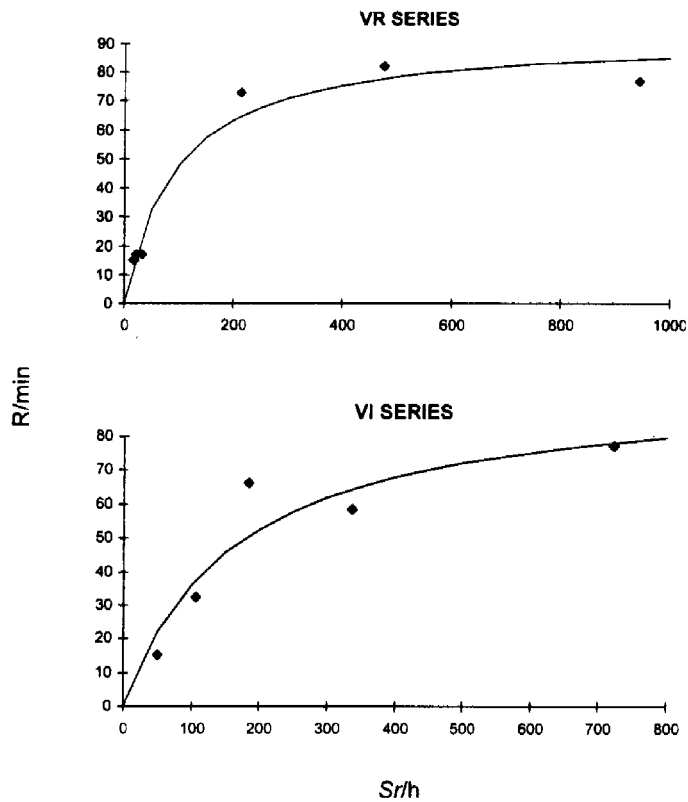


Figure 5. Artie's responses per minute versus obtained reinforcers per hour on a variable-ratio series and a variable-interval series. Both series were fit by hyperbolic functions, resulting in approximately 92.7% and approximately 86.3% of the variance accounted for, respectively

Phase 2 tested the effects of a series of VI schedules on the rate of Artie's lever pressing. Artie was run on schedules VI 100', 200', 500', 1000', and 2000' time steps, where "" indicates that the passage of time was recorded in time steps. Each schedule was in effect for 300,000 time steps. The average rates of lever pressing from the last 10,000 time steps of each schedule was recorded and graphed.

The bottom panel of Figure 5 displays the graph of Artie's parametric VI lever pressing behavior. In the same manner as for the graph in the top panel, the ordinate and abscissa of this graph were calibrated in terms of reinforcers per minute and reinforcers per hour. A fit of a hyperbolic function to the data resulted in approximately 86.3% of the variance being accounted for. Again, this data is similar to data from real animals (see Herrnstein (1970) for data from real organisms).

Experiment 3: Concurrent VI Schedule Series

Experiment 3 tested the ability of Artie's behavior to conform to the matching law when Artie was run on a series of concurrent VI schedules. Operant lever #2 was added to the Skinner box at the previously stated location. Artie was run on concurrent schedules VI 200'-VI 600', VI 400'-VI 600', VI 600'-VI 600', VI 800'-VI 600', and VI 1000'-VI 600', where the "" again indicates that the passage of time was counted in time steps. Before each new schedule, the connection weights in Artie's adaptive network were reset to the values that were saved after Phase 2 of Experiment 1. For each schedule, the reinforcement value was set to 600 and the punishment value was set to -10 for each lever. Each concurrent schedule was run for 300,000 time steps. The response rates and obtained reinforcement rates were recorded from the last 10,000 time steps of each schedule.

Figure 6 presents the results of the concurrent schedule in log-log coordinates. Plotted on the ordinate are the response rates of lever 1 and lever 2 in ratio form, and plotted on the abscissa are the obtained reinforcement rates of lever 1 and lever 2 also in ratio form. A least-squares fit of a straight line through the data revealed a sensitivity value of approximately 1.03, a bias value of approximately 0.85, and a variance accounted for of approximately 96.8%.

DISCUSSION

When subjected to simulated free-operant laboratory experiments, Artie produced behavior that looked like the behavior of real animals. Artie acquired

lever pressing when reinforcement was made available for lever pressing, responded to shaping by learning to lever press at a high rate, ceased responding in response to extinction, and reacquired lever pressing when reinforcement for lever pressing was reinstated. When placed on a series of variable interval and variable ratio schedules, the rate of Artie's lever pressing increased hyperbolically with an increase in the obtained rate of reinforcement. Furthermore, when exposed to a series of concurrent VI schedules of reinforcement, Artie matched his rate of responding on each lever to the relative rate of reinforcement he received from that lever.

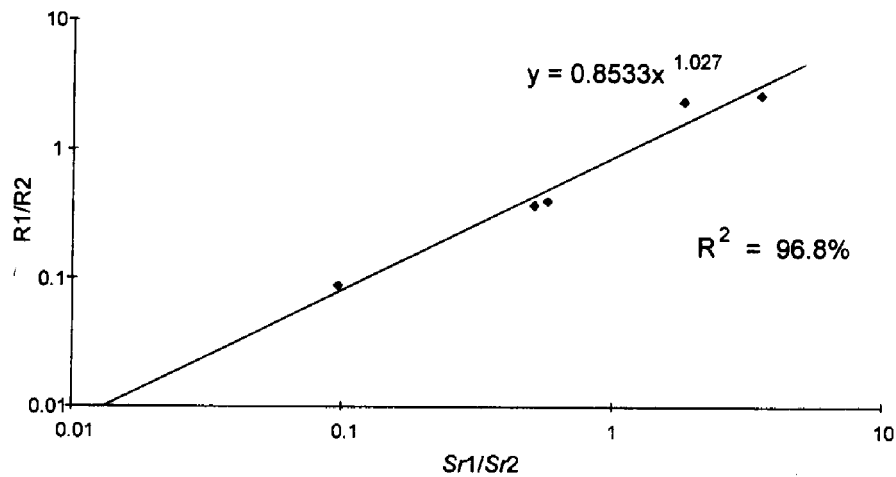


Figure 6. Artie's response rate ratios versus obtained reinforcement rate ratios on concurrent variable-interval schedules of reinforcement, in log-log coordinates

Given that Artie was able to mimic the behavior of real animals in several simulations, what can we learn about real animals from Artie? Can we deduce from the results that real animals are able to adapt to their environments because their brains are conducting essentially the same calculations as Artie's adaptive network? Or should we restrict our inferences to the level of behavior? If it is the latter, then what can Artie, and adaptive networks in general, teach us about behavior?

The utility of adaptive network models of behavior is probably not their ability to teach us about brain function. Marr (1997) has suggested that even though adaptive network models consist of units analogous to biological neurons, the extreme simplicity of adaptive network units in comparison to real neurons limits our ability to extrapolate from one to the other. In addition,

Hutchison (1997) has argued that imposing biological plausibility on behavioral models can limit the model's ability to teach us about behavior.

More likely, the utility of adaptive network models of behavior is their ability to provide convenient methods for testing theories of behavior. All three of the adaptive network models of operant conditioning reviewed earlier showed the sufficiency of selectionist models, and thus of selectionist theories of learning, to account for specific operant phenomena. In addition, two of the models, Stephens and Hutchison (1992) and Spier and McFarland (1998), refuted claims of alternative theories of behavior.

Results from simulations with Artie support selectionist accounts of operant behavior. They also suggest some possibilities for future uses of the Artie model. For example, Artie's concurrent schedule performance was produced purely by "moment-to-moment" changes in environment-behavior relations. If further experiments with Artie can produce matching on a wider range of concurrent schedules, the model could question the necessity of molar accounts of matching. Another use could come from explorations of possible causes of undermatching. Because Artie achieved nearly linear matching, parameters of the model, such as the amount of variation added to behavior, could be investigated for their effects.

REFERENCES

- Beer, R. D. (1990). *Intelligence as adaptive behavior*. California: Academic Press.
- Burgos, J. E. (1997). Evolving artificial neural networks in Pavlovian environments. In J. Donahoe, & V. Packard (Eds.) *Neural-network models of cognition: Biobehavioral foundations. Advances in psychology, Vol. 121*. (pp.58-79). Amsterdam: Netherlands: North Holland/Elsevier Science Publishers.
- Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. *Psychological Review, 67*, 380-400.
- Chomsky, N. (1959). A review of verbal behavior, by B. F. Skinner. *Language, 35*, 26-58.
- Cole, J. L. (1965). The relationship between force gradients and rate gradients in stimulus generalization. *Dissertation Abstracts 25*(8), 4824.
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Needham Heights, Massachusetts: Allyn and Bacon.
- Donahoe, J. W., & Palmer, D. C. (1989). The interpretation of complex human behavior: Some reactions to parallel distributed processing, edited by J. L. McClelland, D. E. Rumelhart, & the PDP Research Group. *Journal of the Experimental Analysis of Behavior, 51*, 399-416.
- Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). A selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior, 60*, 17-40.

- Endler, J. A. (1986). *Natural selection in the wild*. Princeton, New Jersey: Princeton University Press.
- Glenn, S. S., & Madden, G. J. (1995). Units of interaction, evolution, and replication: Organic and behavioral parallels. *The Behavior Analyst, 18*, 237-251.
- Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior, 13*, 243-266.
- Hutchison, W. R. (1984). Cognitive versus behavioral approaches to artificial intelligence. *Newsletter for the Behavioral Artificial Intelligence Network, December*, 5.
- Hutchison, W. R. (1997). We also need complete behavioral models. *Journal of the Experimental Analysis of Behavior, 67*, 224-228.
- Kehoe, J. E. (1989). Connectionist models of conditioning: A tutorial. *Journal of the Experimental Analysis of Behavior, 52*, 427-440.
- Kemp, S. M., & Eckerman, D. A. (1995). Direct analysis of contingencies using working models. *Mexican Journal of Behavior Analysis, 21*, 27-46.
- Langton, C. G. (1986). Studying artificial life with cellular automata. *Physica, 22D*, 120-149.
- Levy, S. (1992). *Artificial life: The quest for a new creation*. New York: Pantheon Books.
- Marr, J. (1997). The eternal antithesis: A commentary on Donahoe, Palmer, and Burgos. *Journal of the Experimental Analysis of Behavior, 67*, 232-235.
- McDowell, J. J., Bass, R., & Kessel, R. (1993). A new understanding of the foundation of linear systems analysis and an extension to nonlinear cases. *Psychological Review, 100*, 407-419.
- McDowell, J. J., & Wixted, J. T. (1988). The linear system theory's account of behavior maintained by variable-ratio schedules. *Journal of the Experimental Analysis of Behavior, 49*, 143-169.
- Moore, J. W., & Choi, J-S. (1997). The TD model of classical conditioning: Response topography and brain implementation. In J. Donahoe, & V. Packard (Eds.) *Neural-network models of cognition: Biobehavioral foundations. Advances in psychology, Vol. 121*, (pp.387-405). Amsterdam: Netherlands: North Holland/Elsevier Science Publishers.
- Notterman, J. M. (1959). Force emission during bar pressing. *Journal of Experimental Psychology, 58*, 341-347.
- Page, S., & Neuringer, A. (1985). Variability is an operant. *Journal of Experimental Psychology, 11*, 429-452.
- Palmer, D. C. (1997). Selectionist constraints on neural networks. In J. Donahoe, & V. Packard (Eds.) *Neural-network models of cognition: Biobehavioral foundations. Advances in psychology, Vol. 121*, (pp.263-282). Amsterdam: Netherlands: North Holland/Elsevier Science Publishers.
- Palmer, D. C., & Donahoe, J. W. (1992). Essentialism and selectionism in cognitive science and behavioral analysis. *American Psychologist, 47*, 1344-1358.
- Plotkin, H. (1994). *Darwin machines and the nature of knowledge*. Cambridge, Massachusetts: Harvard University Press.
- Premack, D. (1961). Predicting instrumental performance from the independent rate of the contingent response. *Journal of Experimental Psychology, 61*, 163-171.

- Reynolds, C. W. (1987). Flocks, herds, and schools: A distributed behavioral model. *Computer Graphics, 21*, 25-34.
- Ridley, M. (1993). *Evolution*. Boston: Blackwell Scientific.
- Rucker, R. (1993). *Artificial life lab*. California: Waite Group Press.
- Shimp, C. P. (1969). Optimum behavior in free operant experiments. *Psychological Review, 76*, 97-112
- Silberberg, A., & Zirix, J. M. (1985). Molecular maximizing characterizes choice on Vaughan's (1981) procedure. *Journal of the Experimental Analysis of Behavior, 43*, 83-96.
- Skinner, B. F. (1986). The phylogeny and ontogeny of behavior. *Science, 153*, 1205-1213.
- Skinner, B. F. (1981). Selection by consequences. *Science, 213*, 501-504
- Smith, T. L. (1983). Skinner's environmentalism: The analogy with natural selection. *Behaviorism, 11*, 133-153
- Smith, T. L. (1986). Biology as allegory: A review of Elliott Sober's the nature of selection. *Journal of the Experimental Analysis of Behavior, 46*, 105-112.
- Spier, E., & McFarland, D. (1998). Learning to do without cognition. In R. Pfeifer, B. Blumberg, J.-A. Meyer, & S. W. Wilson (Eds.), *From animals to animats 5: Proceedings of the fifth international conference on simulation of adaptive behavior* (pp.38-47). Cambridge, Massachusetts: MIT Press.
- Stephens, K. R., & Hutchison, W. R. (1993). Behavior analysis and intelligent machines. *Educational Technology, 23*, 52-61.
- Stephens, K. R., & Hutchison, W. R. (1992). Behavioral personal digital assistants: The seventh generation of computing. *The Analysis of Verbal Behavior, 10*, 149-156.
- Stokes, P. D. (1995). Learned variability. *Animal Learning & Behavior, 23*, 164-176.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge, Massachusetts: MIT Press.
- Terzopoulos, D., Tu, X., & Grzeszczyck, R. (1994). Artificial fishes with autonomous locomotion, perception, behavior, and learning in a simulated physical world. *Artificial Life, 4*, 17-27.
- Weiner, J. (1994). *The beak of the finch*. New York: Random House.