GENERALIZATION OF LEARNING IN A MACHINE

R. J. Lee
Melpar, Inc., Falls Church, Virginia

An artificial cerebral cortex which incorporates machine learning in general form is described. The scheme offers promise for new lines of basic research in self-programming data processing.

A small mobile vehicle (suggestive of an insect) that can wander about in a suitable environment, directed by the cortex model, will illustrate the system concept in simple terms. The vehicle will feed information into the cortex via two photoelectric "eyes," a microphone "ear" sensitive to several audio bands, and two "feelers" that point forward and to the sides. Its outputs will be to turn left, turn right, or head forward. The environment will contain a battery charger, various obstacle configurations, and patterns of sound and light.

Interacting directly with the environment, the machine will learn a wide range of uncomplicated routines for satisfying two motivations, which are

(1) keep an internal storage battery charged
(2) continue in motion, avoid obstructions.

These motivations are effected by circuits outside the cortex which feed the cortex. In order to satisfy its motivations, the machine can, for example, learn to navigate a curving corridor without becoming immobilized against a wall, learn to avoid obstacles, solve a T-maze with the charger in one branch, and relearn the maze if it is reversed, or seek a lighted or a darkened target marking the charger.

The artificial cortex is not tailored to any of these activities, but rather is set up to construct spontaneously its own program for solving whatever problems may stand in the way of satisfaction of its motivations. The cerebral cortex is thus emancipated from limitations imposed by a designer or programmer having to know ahead of time the nature of the problems to be solved. This results in generalization of learning ability.

For such an approach, it is appropriate to consider simple problems at this time, but to study intensively the learning or conditioning process.

The cortex consists of similar decision units, which are comparable to living neurons. The units are called <u>artrons</u>, a contraction of <u>artificial</u> <u>neurons</u>. An artron has two input channels <u>a</u> and <u>b</u> as <u>artificial dendrites</u>, one branching output channel <u>c</u> as an artificial axon, and two input channels R and P which may correspond to chemical influences from the blood stream.

The artrons are stratified in layers and are randomly soldered to connect axons of one layer to dendrites of the next layer or to dendrites of a previous layer. Sensory channels connect to dendrites of the first layer at random and axons of the final layer connect to effector channels at random. R and P denote reward and punishment. These two channels are pathways for outputs of the motivation center. They exert statistical control over the logical transfer-functions, or states, of the artrons. All R input channels are in parallel and all P input channels are in parallel.

Seven such states will be treated here. They are: repeater on a ignoring b, repeater on b ignoring a, and gate, repeater on a inhibited through b, repeater on b inhibited through a, or gate, and open circuit. The and and inhibit states are time-dependent. In the and state, positive probability of output does not require coincidence of input signals if the two signals (on channels a and b) are close together. Also, inhibitory effects may endure longer than the inhibitory signal.

In the unconditioned cortex, artrons may adopt any one of these seven states, and randomly fluctuate from one to the other, subject to the following constraints:

(1) each artron will average a comparable portion of time in each state

(2) changes of state occur together in periodic beats, separated by equal intervals of no change, it being uncertain whether or not a particular artron will change state on a particular beat.

Learning involves statistical modification of the persistence of each state in each artron through time-correlation with the overall success of the machine in terms of the two motivations. When the situation improves for the machine (by the battery receiving charge or by the vehicle resuming motion after having been blocked) the R buss is automatically pulsed. When matters deteriorate (from falling battery charge or from obstruction of the vehicle) the P buss is automatically pulsed. Pulses on the R and P channels influence states in effect at the time or at a shortly previous time in artrons which have concurrently received signals on the a or b channels. An R pulse acts to increase the persistence of such a state in a particular artron; a P pulse acts to decrease persistence.

Chains and networks of artrons may now evolve and stabilize through a process suggestive of natural selection. For example, artron chains from the left feeler channel to the right-turn-command channel and from the right feeler channel to the left-turn-command channel will guide the vehicle down a corridor. Resumption of motion after turning away from the corridor wall provides the reward. A third chain carrying one feeler signal to inhibit the opposite feeler chain will produce a 90-degree turn in a T-maze and guide the vehicle around obstacles. Chains from the "eye" channels to the command channels for turning can provide target-seeking capability. Furthermore, any required logic transfer-function may arise and be stabilized, in a large enough network. Prior to learning, typical artrons were undifferentiated with respect to sensory and effector channels.

Although an artron-type machine can learn a wide range of elementary problems, given sufficient time, it is desirable to know statistically the amount of time required, and what can be done to improve learning-speed for such problems. These questions cover the primary mathematical area of effort at present.

Approximate learning charts are derived for a very simple case involving nine input channels, under optimized conditions, as a first step in constructing the theory of artron networks. This involves learning-time for establishing a chain of artrons from a given (sensory) input channel to a given (effector) output channel without having the chain acted upon by signals from any one of the eight other input channels. It also involves treatment of conditioning effects that shift the machine's behavior from random toward determinant in a series of increasingly frequent successes. Only R pulses are considered in this simple case.

For the vehicular system, where average state lifetimes prior to learning are adjusted to performance limitations imposed by traction, inertia, speed, and other mechanical properties, conclusions are reached that:

(1) Average time for setting up the first satisfactory chain will be slightly less than thirty minutes.

(2) Though learning speed is adjustable by varying the effect of the R pulses, it is not desirable to produce complete fixation of states with the first success. After only one reward, a chain will break up, but will be favored to reoccur.

(3) With reasonable adjustments of one learning parameter, the machine will cease to be noticeably "confused" or to make an appreciable number of mistakes, in one or two hours. With fewer than nine sensory inputs stimulated, learning will be faster.

(4) Through learning, not just one, but numerous chains of the required type will be established in parallel.

As an example, the artron chain affected by only one sensory channel is important in teaching the equipment to turn right when "hearing" a particular tone and left when "hearing" another particular tone. This will involve evolution of a chain from one audio filter channel to the command channel for steering left and a chain from a second audio filter channel to the command channel for steering right. Both the left action and the right action may be learned concurrently, so the total learning-time will be one or two hours. A human trainer may be employed here. Between the phased artron changes, he should provide first one tone and then the other. He should reward proper response by administering a brief battery charging through clip leads.

Learning can be speeded up many orders of magnitude if incoming information rates are not restricted by limitations of the vehicle. The vehicle is simply a useful theoretical device for studying and for illustrating the artron system.

The author first developed the basic qualitative concepts of the artron network in 1952. He presented them at a series of group meetings in 1953 at the University of Oklahoma. More recently he has developed the probability description of the system, which will constitute a major portion of the full presentation.