A Possible Biological Analogue of the Reinforcement Control Device in Self-Organizing Systems

A. A. Buerger

Department of Psychology, Harvard University, Cambridge, Mass.

Abstract—Virtually all self-organizing systems proposed as models for learning require a reinforcement control device. The biological analogue of this device has never been clear. Slow changes in the cortical potential may be correlates of a biological analogue of reinforcement control devices.

Frequent pairings of novel stimuli with other stimuli eliciting specific responses produce a slow shift and oscillation in the cortical potential. This D.C. shift and oscillation occurs first in the sensory, and then in the motor cortex involved; it may be associated with the decreases or increases which occur in the threshold of cells in these cortical areas.

A self-organizing system may be described loosely as any system which modifies its behavior in accordance with inputs from the environment in which it operates. This modification of behavior is accomplished by the system itself, without external intervention in the functioning of its components. If a self-organizing system is successful, the result of the behavioral modification is improved performance according to some criterion—usually a reduction in errors or improved stability.

Self-organizing systems have many features which make them attractive models for mammalian learning. There are five basic components which are found implicitly in these models. They include the environment, sensors, neuromimes1 response indicators, and a "trainer" or reinforcement control device. The biological analogues of four of these components are obvious. The environment of the system is clearly analogous to the environment of an organism, the sensors to various sensory receptors, the neuromimes to neurons, and the response indicators to muscles, glands, etc.

However, the biological analogue of the reinforcement control device is not clear. This is a serious problem because virtually all self-organizing systems proposed as models for learning require such a component (Hebb, 1949; Ashby, 1952; Audley and Jonckheere, 1956; MacKay, 1956; Milner, 1957; Minsky, 1961; Rosenblatt, 1962).

1. Neuromimes, as defined by van Bergeyk (1960), are man-made devices which mimic some or all of the properties of nerve cells.
Low-Frequency Long-Lasting Potentials as a Correlate

The study of the electrical activity of the brain during classical conditioning and during other learning situations may have begun to identify such an analogue. Russian studies of low-frequency, long-lasting potentials during cortical conditioning are especially suggestive. (The relevant Russian papers are now available in English and will be cited in translation.)

Many American and British papers deal with the origins of low-frequency, long-lasting potentials (Brookhart, Arduini, Manzia and Moruzzi, 1958; Caspers, 1961; Gumnit, 1961) and their significance during conditioning (Morell, 1960, 1961; Rowland and Goldstone, 1963). None are directly relevant here, although the results of Morrell (1960, 1961) and of Rowland and Goldstone (1963) sometimes support the conclusions suggested in this study. Three worthwhile reviews of this literature have been published recently (Rowland, 1967, 1968; Adey, 1969). This report will not attempt to review the literature relating slow potentials to various types of adaptive behavior. Rather, it will attempt to prove that many studies strongly suggest that slow potentials may be correlates of reinforcement control devices in the central nervous system.

Low-frequency, long-lasting shifts in cortical potentials accompanying higher nervous processes were first reported by Kohler et al. (Köhler, Neff and Wedner, 1955; Köhler and O'Connell, 1957). These shifts were not studied as a function of classical conditioning but were suggested to correlate with learning not to respond to inconsequential stimuli.

Shvets (1960; Rusinov, 1961), an associate of Rusinov, studied low-frequency, long-lasting (henceforth called slow, or D.C.) potentials of the frog's and rabbit's cerebral cortex during the formation of conditioned reflexes. In the frog experiments, light flashes were paired with shocks to the left foreleg. Such pairings induced gradual shifts in the slow potential of the cerebral cortex with durations of approximately 30 seconds. Neither light nor shock applied alone produced these shifts in the slow potential. Similar results were observed in rabbits (Fig. 1).

Early in the formation of the conditioned reflex, when the light was just beginning to "cause" withdrawal of the foreleg, slow and sustained oscillations of the D.C. potential were recorded at the visual cortex. These oscillations had latencies of one-seven seconds, amplitudes of 200 μv to several millivolts, and durations of five-50 seconds. After more pairings of light with shock, these oscillations in the D.C. potential of the visual cortex were intensified; after still more