

NEURAL COMPUTATION

[In this five-part article, Approaches to Learning—an overview of biological neural networks and their relationship to artificial neural networks—is followed by examinations of neural computation as it is evidenced in the functioning of the cerebellum, the hippocampus, the neocortex, and the olfactory cortex.]

Approaches to Learning

Nervous systems are capable of solving extraordinarily sophisticated computational problems. The visual or tactile recognition of an object in a cluttered scene is child's play, but well beyond the capability of the fastest digital computers. Most animals can navigate over rough surfaces with great agility, but present-day robots are limited in their movements to a very narrow range of terrains. We can learn to use language and to read and write, well beyond anything so far accomplished by artificial intelligence. We take all of these abilities for granted because we are so good at them; trying to duplicate them with machines has made their great difficulty more apparent.

Neural computation is the systematic study of the computational principles underlying the function of neural systems, from the level of molecular mechanisms to the organization of brain systems (Figure 1). This computational approach to neuroscience is still in its infancy (Sejnowski et al., 1988). There has been a recent emphasis on studying neural networks, small groups of highly connected neurons; however, as shown in Figure 1, neural networks are only one level of investigation in the nervous system, and neural computation depends on computational principles at each of these levels. A few general principles have emerged from the study of abstract models of neural systems

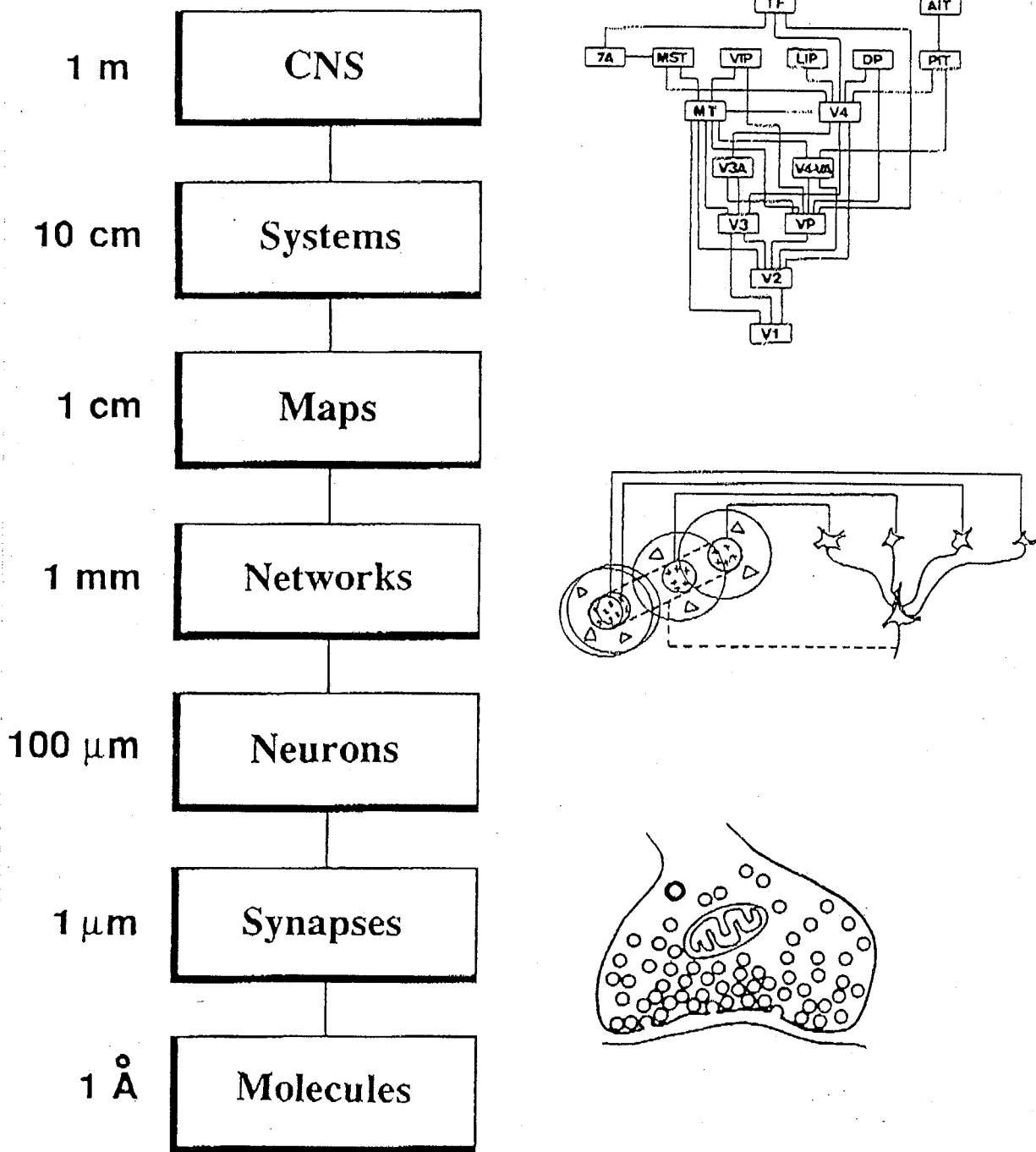


Figure 1. Structural levels of organization in the nervous system. The spatial scale at which anatomical organizations can be identified varies over many orders of magnitude.

that are likely to be important for the biological study of learning and memory.

Some Principles of Neural Computation

In the von Neumann architecture commonly used in digital computers, the memory and the processor are physically separated. This separation gives rise to a bottleneck in the flow of information between the two. In neural systems, memory and processing are intertwined; the same circuits that process sensory and motor information are involved in learning and the storage of new information. A unified processor-memory system allows many circuits to work together in parallel and, as a consequence, the solutions to many commonly occurring problems can be computed in only a few serial steps. The representation of sensations and memories in such an architecture is more difficult for us to imagine and to use than one in which the functions are segregated (Churchland and Sejnowski, 1992). The brain, however, did not evolve to make it easier for us to analyze.

Locality is an important constraint that arises when hardware for artificial neural networks is designed (Mead, 1989). Wires are expensive on computer chips, just as they are in the brain, so only limited connectivity is possible between processing elements. The organization of sensory processing into a hierarchy of maps and the laminar organization of cortical structures is wire-efficient. This also places constraints on the organization of learning systems, which share the same circuitry. In particular, the decision to store a piece of information at a particular location in the brain is a local one that depends on electrical and chemical signals which are spatially and temporally restricted. The Hebbian mechanism (see HEBB, DONALD) for synaptic plasticity that has been found in the hippocampus and neocortex obeys this principle of locality: The presynaptic release of neurotransmitter and the postsynaptic depolarization needed to trigger the long-term potentiation at these synapses are spatially contiguous, and there is a brief temporal window which both signals must be present. Modulatory influences on learning may be more diffuse and widespread.

Neurons have limited dynamic range. Unlike digital systems, which are capable of accurately representing very large and very small numbers, the range of membrane potentials and firing rates

found in neurons is limited. Also, the variability in the properties of neurons within the same population is significant, and the properties of the same neuron can vary with time. The same is true for analog VLSI (very-large-scale integration) circuits that are designed to mimic the processing which occurs in neurons (Mead, 1989). This variability and limited dynamical range have consequences for the way that information is coded and the way that neural circuits are designed. One way to preserve information is to process relative levels, or differences, rather than absolute levels. Thus, visual neurons are more sensitive to contrast (spatial differences) and changes (temporal differences) than to absolute intensity levels. Another mechanism for preserving information is dynamically altering baseline activity in neurons. Adaptive biochemical mechanisms inside cells, such as light adaptation in photoreceptors, allow neurons to remain in their most sensitive range. Adaptive mechanisms have been found for calibrating sensorimotor coordination, such as slow adaptation of the vestibulo-ocular reflex (VOR) to changes in the magnification of the lens (Lisberger, 1988).

Taxonomy of Learning Systems

Adaptation to ongoing sensory stimulation does not require an additional source of information outside the processing stream; this type of learning is called *unsupervised*. In contrast, the type of adaptation that occurs in response to sensorimotor mismatch does require an outside error signal; this is called *supervised* learning. In the case of VOR learning, the error signal is the slip of the image on the retina, and the gain of the reflex is changed to reduce the slip. The amount of supervision can vary from a crude good/bad reinforcement signal to very detailed feedback of information about complex sensory signals from the environment that might be termed a "teacher." Supervised learning is sometimes called *error-correction learning*.

It is not necessary for the error signal to come from outside the organism; important information about the proper operation of a circuit can be provided by another internal circuit, or even by internal consistency within the same circuit. For example, a sensory area that was trying to predict future inputs could compare its prediction against the next input to improve its performance. Such an unsupervised system with an internal measure

of error is termed *monitored* (Churchland and Sejnowski, 1992). As shown in Figure 2, all possible combinations of supervised and unsupervised, monitored and unmonitored, learning systems are possible.

Selected Examples

An interesting example of a monitored system is song learning in the white-crowned sparrow. In this species of songbird, the male hears the local dialect after hatching but does not produce the song until the next spring. At first the song is imperfect, but with each repetition the details improve until it is a good reproduction of the original song heard the previous year, even though there is no external "teacher" during the refinement. The internally stored template is compared with the imperfect song production; the error between them drives learning mechanisms to improve the song. This learning is monitored because the error is derived from an internal template of the desired sound. We may use a similar strategy when learning to produce new sounds in a foreign language. (See also BIRD SONG LEARNING; LANGUAGE LEARNING.)

Transformations between two populations of neurons can be learned with Hebbian mechanisms at the synapses between the input fibers and the output neurons. The pattern of activity on the input fibers is matched with the desired pattern on the output neurons. In some models of the cerebellum, for example, associative motor learning is mediated by climbing fibers, which provide a teaching signal to the output Purkinje cells. By including feedback

projections of the output neurons back onto themselves, a partial input cue can regenerate the entire stored pattern. In this mode, the system is unsupervised because the desired output pattern of activity during learning is the same as the input pattern. Such content-addressable recurrent networks have been suggested as models for the piriform cortex and area CA3 of the hippocampus. Some properties such as memory capacity of associative networks of simplified processing units have been well studied; the analysis of networks based on model neurons with more complex properties is just beginning.

Learning mechanisms have also been used to model the development of cortical systems. One of the best-explored areas of unsupervised learning in artificial networks is competitive learning, in which incoming sensory information is used to organize the internal connections of a sensory map. For example, the formation of ocular dominance columns in visual cortex of cats and monkeys during development (see NEOCORTICAL PLASTICITY) depends on competitive synaptic mechanisms. The development of ocular dominance columns can be mimicked in a computer model that uses Hebbian learning in the spatially restricted terminal arbors of axons projecting to the cortex from the lateral geniculate nucleus (Miller et al., 1989). Similar mechanisms can also be used in neural systems to learn more complex features that distinguish among different types of sensory inputs (Kohonen, 1984). It is also likely that other learning mechanisms are used to discover invariants of sensory patterns, which are often as important in pattern recognition as the distinctive features that separate classes (Churchland and Sejnowski, 1992).

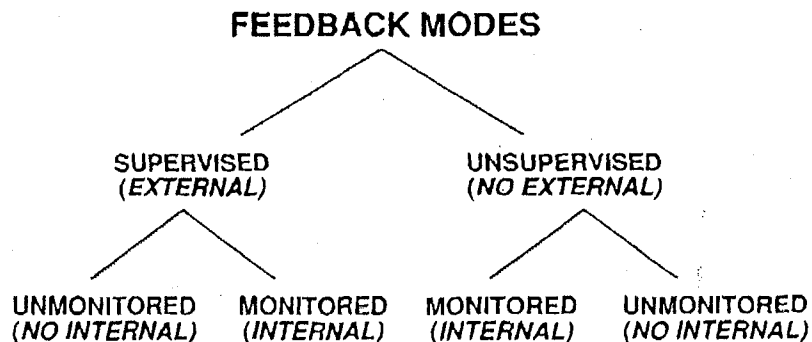


Figure 2. Taxonomy of learning procedures. Supervised learning occurs when there is feedback on the performance of the system from the external environment. If the feedback is a scalar reward, it is called reinforcement learning. The learning is called monitored if the system has an internal measure of error.

As computers grow more powerful, it will become possible to simulate more complex models of neural systems; however, even these simulations will fall short of the richness of real neural systems and the complex environments that confront biological creatures. Hardware emulations that interact with the real world in real time would greatly improve our ability to test hypotheses about the organization of neural systems (Mead, 1989). Ultimately we will need to study complete model systems in order to understand the multiple levels of adaptation and learning that provide flexibility and stability in a changing world.

REFERENCES

- Churchland, P. S., and Sejnowski, T. J. (1992). *The computational brain*. Cambridge, Mass.: MIT Press.
- Kohonen, T. (1984). *Self-organization and associative memory*. Berlin: Springer-Verlag.
- Lisberger, S. G. (1988). The neural basis for motor learning in the vestibulo-ocular reflex in monkeys. *Science* 242, 728.
- Mead, C. (1989). *Analog VLSI and neural systems*. Reading, Mass.: Addison-Wesley.
- Miller, K. D., Keller, J. B., and Stryker, M. P. (1989). Ocular dominance column development: Analysis and simulation. *Science* 245, 605-615.
- Sejnowski, T. J., Koch, C., and Churchland, P. S. (1988). Computational neuroscience. *Science* 241, 1299-1306.

Terrence J. Sejnowski