CHAPTER 2

NERVOUS CONDUCTION: EARLY CONNECTIONISM, PHYSICAL ANALOGIES AND MACHINES

1. Introduction. Neurological hypotheses and machines

In the previous chapter we saw that for Loeb associative memory was a mark of intelligence and learning capacity. His thesis was that the lower organisms did not have associative memory, a thesis that seemed to be corroborated by a machine that was actually built by an engineer, Benjamin Miessner. The machine displayed behavior that was indistinguishable from that of a lower organism, and since it was a machine, and since it was automatic, it was by definition without associative memory or the ability to learn.

At the time, another engineer, S. Bent Russell, was discussing the possibility of a machine endowed with memory that could simulate the working of a nervous system capable of rudimentary forms of learning. Paradoxically he appealed to the authority of Loeb, in part to explain the importance of his own machine (had it not been Loeb who considered the function of associative memory the "great discovery" to be made in physiology and psychology?) and in part to identify the ability to learn as the criterion for showing that the machine itself had associative memory (had it not been Loeb who said that "if an animal [...] can learn, it possesses associative memory"?) (Russell, 1913: 20 and 35). With a misunderstanding not unlike that of Miessner, who credited Loeb with an anthropomorphic explanation of reflex action, Russell interpreted in terms of *inorganic* machines associative memory, which Loeb considered a proper subject for physiology, not engineering, a feature of some organisms but never of machines.

Russell saw no particular difficulty in the physical explanation of the nervous discharge that occurs in simple reflex action, where a given stimulus always evokes the same response. So he proposed to simulate with his mechanical device some phenomena of conductivity in the nervous system that he called "perplexing": they were phenomena in which the same stimulus could evoke different responses or in which different stimuli evoked the same response, namely phenomena displayed by organisms capable of learning.

The following is Russell's simulation method, as expounded in his 1913 paper in the *Journal of Animal Behavior*. He first summarized the hypotheses on nervous conduction that seemed to be generally accepted by psychologists, then described a hydraulic machine that he said "embodied" these hypotheses, and proceeded to "compare the results obtained with the machine with those given by live nervous connections." This comparison made him think it possible that a mechanical device could simulate the "essential elements" that, according to the most widely accepted neurological hypotheses, accompanied inhibition, habit formation and other forms of learning. He concluded with a study of the differences between machine behavior and that of the nervous system, expressing the hope that "with some modification" the machine would be able to simulate more complex forms of learning.

It was a method that seems quite sophisticated for the period in which it was proposed, exactly thirty years before 1943, the year which is usually considered to mark the birth of cybernetics. With its potential and its limitations this method marked an important point in the history of the sciences of the mind and behavior. In the present chapter and the next two we shall follow some of the germinal phases only hinted at in Chapter 1.

But which hypotheses on nervous conduction inspired Russell when he designed his machine? The authors he mentioned were Herbert Spencer, Max F. Meyer, and Edward L. Thorndike. These are fairly disparate and heterogeneous figures, but they all shared the associationist slant and the idea that the variability of behavior observed in habit formation and learning might be explained as the result of some modification of nerve tissue. Anyway, the authors Russell mentioned disagreed on the exact nature of that modification. In the wake of Spencer and Meyer, there emerged several versions of the so-called "drainage theory" of nervous conduction, a theory that James picked up from Spencer and that was refined by McDougall. The analogy that inspired the theory was that of nervous current as continuous fluid flowing through a network of natural channels or pipes connected by one-directional valves. These valves were later made to correspond to synapses. The deviation of a nervous current from one direction to another, which would explain learning phenomena, occurred through a physical process comparable to that of a liquid or fluid being drained. As for Thorndike, he had proposed a theory of nervous conduction as an impulse transmitted by a complex system of "connections" or "bonds" between stimulus and response (S-R). These connections could be strengthened or weakened depending on the effect produced by the response, which could subsequently modify the direction of the nervous current and give rise to learning. For some traits at least, this system was compared to a telephone network of countless units, corresponding to neurons, that allowed the nerve message to be routed through an enormous multiplicity of conduction lines.

In the present chapter we shall first examine the different concepts of learning that go back to these two theories of nervous conduction. Thus we can make clear which neurological hypotheses Russell embodied in his machine and, most of all, what was new in his approach as compared with the mechanical analogies used not only by mechanistic psychologists like Meyer but also by vitalist psychologists and biologists like McDougall and von Uexküll. Russell's machine was designed as a working mechanical model of some "perplexing" phenomena of the nervous system and worked according to certain neurological hypotheses embodied in it, but the hydraulic analogies for the nervous system, based on drainage, came down to partial and forced images of fluid dynamics not infrequently bolstered by ad hoc explanations. Furthermore, the use of these analogies depended on whether or not a purely physical explanation of learning and mental life seemed sufficient, and this gave rise to the conflict between "Automaton Theories" and "Ghost Theories," which is how McDougall and Meyer, respectively, polemically characterized those theories that refused to introduce a non-physical cause or agency into the explanation of learning and mental life and those theories that instead considered it indispensable. We shall mention the later evolution of Russell's neurological thinking, which led him to support hypotheses about nervous conduction that seemed equally distant from drainage theory and from Thorndike's theory and altogether closer to certain of Watson's claims. We shall end the chapter by describing a simulative methodology even more explicit and resolute than Russell's, proposed by the psychologist J. M. Stephens. He tried to embody in a simple electro-mechanical device the law of effect, the most important of the laws of what Thorndike called the "connectionist" theory of learning.

However rudimentary Russell's machine and Stephens' were, they are a novelty in the field of the study of organism behavior when they are compared with the aforementioned scanty mechanical analogies. Notwithstanding, we shall dwell on these analogies at length in this chapter. On the one hand, they constitute a little-known issue in the history of psychology and neurology and stimulated the rise of a connectionist view of nervous activity. On the other hand, the very intricate difficulties in which those analogies were involved show up the actual novelty of a modeling approach, which lies in building working artifacts that embody explanatory hypotheses on behavior, and as such is at the core of the discovery of the artificial.

2. Nervous conduction as drainage: Spencer, James and McDougall

Aside from vague references by British empiricists, David Hartley in particular, it was Spencer who explicitly connected the formation of mental associations with the formation of nervous connections.[1] In the following, almost forgotten passage of the first edition of his *Principles of Psychology*, Spencer formulated a law of formation and "consolidation," as he put it, of the connections between parts or nervous states of the organism, a law with which modern neurological associationism made its debut in the wake of psychological associationism.

> The law in virtue of which all physical states that occur together tend to cohere, and cohere the more the more they are repeated together, until they become indissoluble [...] is the law in virtue of which nervous connections are formed. When a change in one part of an organism is habitually followed by change in another; and when the electrical disturbance thus produced is habitually followed by a change in another; the frequent restoration of electrical equilibrium between these two parts, being always effected through the same route, may tend to establish a permanent line of conduction—a nerve (Spencer, 1855: 544 n.).

So what Spencer called the "general law of intelligence," which explained the association of ideas, was the same law that should explain the formation of new nerve connections. These gave rise not to "perfectly automatic" responses, like those aroused by reflex activity and instinct, but to unpredictable and initially "irregular" responses, as Spencer put it, which showed that memory must be present. And this was what made it possible for the organism to adapt to the environment in a not wholly passive way, that is to say, to acquire new habits in the form of *new* automatisms and to learn.[2]

The description of the hypothetical physical mechanism of the *formation* of nerve connections underlying association was very vague, and that continued to be the case in sub-

[1] James (1890) mentioned these antecedents, and the story was subsequently reconstructed by McDougall (1911, Chapter 8).

[2] See above, p. 11.

sequent editions of the *Principles*, which were extensively rethought and expanded. Although this hypothetical mechanism was later described in different terms, Spencer continued to think of an organism's nervous system as a system of forces in equilibrium. In general, every sensory stimulus from the external environment was treated as an alteration of that equilibrium, which could be followed by its restoration. According to Spencer, when the organism was stimulated from outside, it was as if a force were added or liberated at a particular point A of its nervous system, which force, being resisted by smaller forces around, gave rise to a motion towards other points in the nervous system. Suppose that there was a point B elsewhere in the organism at which a force was being expended. In this case motion would arise between A and B, and thus in the nervous system a "line of least resistance" would have been forced, like a drainage channel, between point A, at which excess energy was liberated and point B, where it could be absorbed. In this way equilibrium would be restored in the nervous system (Spencer, 1855/1890, vol. 1: 511-512). This, for example, is the way the phenomenon of a stimulus in a specific part of the organism habitually followed by a contraction in some other specific part was treated.

Figure 2.1a shows Spencer's "network of lines of least resistance" in a nervous system where one of the "afferent fibers" A runs from point *a* to all points *e* of the "efferent fibers" E. If tension is established between point *a* and all points *e*, it must be susceptible to suppression by flows through the connections between point *a* and points *e* for the system to return to equilibrium.

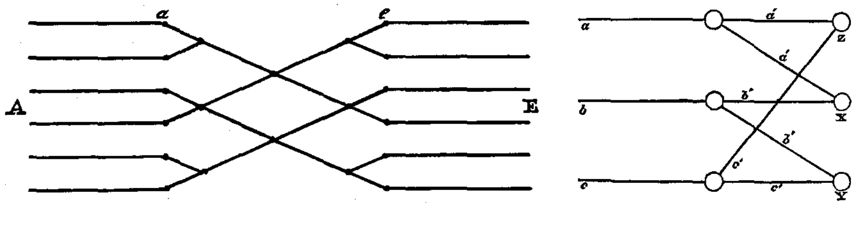
Spencer distinguished the formation of new connections between different points or centers of the nervous system from the *facilitation* or *fixation* of connections once they are formed. Facilitation consisted in the gradual increase of permeability of the connections, and it was repeated use that turned a "vague course," the newly formed connection, into a "definite channel" with a lower resistance, in which the passage of the nervous impulse could gradually be made easier (p. 515). In this way a "permanent line of conduction" is fixed according to the above-mentioned general law of intelligence, and thus the organism acquires the ability to make a new association; actually he becomes able to change his response and to learn.

A similar description of the formation and fixation of nerve connections can be found in Bain's *Mind and Body*.[3] Bain's neurological statement of the law of association is as follows:

> In consequence of two [nervous circuits] being independently made active at the same moment, [...] a strengthened connection or diminished obstruction would arise between these two, by a change wrought in the intervening cell-substance (Bain, 1873: 119).

Figure 2.1a shows Bain's "hypothetical scheme" of nervous connections (p. 111). The unequal intensity of stimulation of fibres *a*, *b*, *c* causes nervous currents of different intensity, and this difference of intensity ends in a difference in the nervous circuits converging in cells *X*, *Y*, *Z*, which are supposed to be the commencement of motor fibres, each concluding in a distinctive movement or response. Once new connections or nervous cir-

[3] I am indebted to Carmela Morabito for drawing my attention to this point.



(a) (b)

Fig. 2.1. Networks of nervous connections: (a) from Spencer (1855/1890); (a) from Bain (1873).

cuits are formed, they are to become fixed, in order to have "the physical bond underlying memory, recollection, or the retentive power of the mind", as Bain put it (p. 116)

Bain's and Spencer's descriptions of the formation and fixation of a nervous connection are rather obscure. Bain quite briefly used the analogy of electrical induction in order to give an idea of the diffusion of nervous current. As to Spencer, he recurred to the aforementioned analogy of drainage, but it is not sufficiently clear how drainage actually works when different cortical centers are repeatedly stimulated in immediate succession, the physical correlate of the association of ideas. Among psychologists, it was James in his *Principles of Psychology* who tried to formulate more detailed hypotheses about the nature of this phenomenon.[4]

James thought that, generally speaking, for any physical system to form new habits and maintain them through the course of time, it must be capable of changing its internal structure or organization as well as putting up some "resistance" to the cause for change. Only in this way could a physical system achieve a new internal structure or organization without so to speak disintegrating, i.e. it would be capable of reaching a new state of equilibrium. This kind of system must necessarily be *plastic*. "Plasticity [...] in the wide sense of the word," James wrote, "means the possession of a structure weak enough to yield to an influence, but strong enough not to yield all at once. Each relatively stable phase of equilibrium in such a structure is marked by what we may call a new set of habits" (James, 1890, vol. 1: 105).

There has always been interest in the existence in the physical world—the world of "dead matter," as James called it—of this kind of plasticity, whereby a system's behavior can be influenced by its previous history. We have already run into the observations of Loeb and Jennings about 'memory' found in the inorganic world in the form of different kinds of hysteresis, and Spencer also commented on the physical analogies that might be made with some phenomena of organic memory. James too was rather attracted by the superficial aspect of these analogies, so much so as to make comparisons between water flow and nerve currents, between a natural channel's resistance to water and that of nerve paths: "wa-

[4] The literature on James' *Principles* is vast. Suffice it to mention the collection of essays edited by Johnson and Henley (1990) on the centenary of the book's publication. Walker (1990) deals with James (and Spencer) in the history of connectionism (see Chapter 6 below).

http://www.springer.com/978-1-4020-0606-7

The Discovery of the Artificial
Behavior, Mind and Machines Before and Beyond Cybernetics
Cordeschi, R.
2002, XX, 314 p., Hardcover
ISBN: 978-1-4020-0606-7