A Theory of the Epigenesis of Neuronal Networks by Selective Stabilization of Synapses

(control theory/graph theory/learning/synaptic plasticity/neuromuscular junction)

JEAN-PIERRE CHANGEUX, PHILIPPE COURRÈGE*, AND ANTOINE DANCHIN

Institut Pasteur, 25 rue du Dr. Roux, 75015 Paris; and * Institut Henri Poincaré, 11 rue Pierre et Marie Curie, 75231 Paris

Communicated by Jacques Monod, July 3, 1973

ABSTRACT A formalism is introduced to represent the connective organization of an evolving neuronal network and the effects of environment on this organization by stabilization or degeneration of labile synapses associated with functioning. Learning, or the acquisition of an associative property, is related to a characteristic variability of the connective organization: the interaction of the environment with the genetic program is printed as a particular pattern of such organization through neuronal functioning. An application of the theory to the development of the neuromuscular junction is proposed and the basic selective aspect of learning emphasized.

We relate the epigenetic development of the nervous system to learning and define learning as the process by which a complex organism acquires a well-defined and stable associative property as a result of a specific interaction with environment. Such a process and the resulting property can be considered at two distinct levels: the *functioning* of the neuronal network (electrical activity) and the *behavior* of the entire system constituted by the neuronal network plus its relevant environment in reciprocal interaction. In this paper, after formulating postulates (1) we shall propose a class of mathematical models† that represent the joint structural evolution and functioning of a neuronal network (*Section 1*), the behavior of the organized system associated with it (*Section 2*), and, finally, the relation between the two levels of learning through semantics of the genetic program (*Section 3*). In *Section 4* the theory is applied to development of the neuromuscular junction.

Biological Postulates. (P1) The physical basis of the informative operations performed by the nervous system resides in the ability of the nerve cells (or neurons) and their axonic or dendritic processes to produce, transmit, integrate, propagate ... impulses through the complex neuronal network. All main features of that functioning may be registered by an absolute time scale. (P2) The interneuronal contacts, the synapses, mediate all information transfers through the system. Excitatory as well as inhibitory synapses may exist under at least three connective states (synaptic "plasticity"): labile (L), stable (S), and degenerate (D); only states L and

S transmit nerve impulses and the acceptable transitions between states are $L \to S$, $L \to D$, and $S \to L$. (P3) Evolution of the connective state of a given synapse is governed by the total message afferent to the postsynaptic soma during a prior time interval of determinate length (evolutive power of the soma). (P4) The maximum wiring and the main stages of development of the network of synaptic connections, as well as the evolutive power (postulate P3) and the integrative power (after the usual "firing" mechanism) of each soma is a determinate expression of the genetic program (the "genetic envelope" of the network). The emergence during growth of a large number of labile synapses is provided by this program. (P5) Neuronal learning appears as a capacity associated with the variability of the connective organization of the neuronal network. The associative property that results from the learning process is structurally printed as a particular pattern of such organization; this pattern results often from the selection by functioning (postulates P2 and P3) of particular pathways among a large number of labile synapses (especially during growth; postulate P4).

1. Statics and dynamics of neuronal networks

1.1 Neuronal Graphs and Neuronal Networks. If C is a finite set, and if $\Sigma \subset C \times C$, we set $\Sigma(C) = \{x \in C \mid \exists \, y \in C, (y,x) \in \Sigma\}$, $\Sigma^{-1} = \{(y,x) \mid x \in C, y \in C \, \& \, (x,y) \in \Sigma\}$, $A(C,\Sigma) = C \diagdown \Sigma(C)$. $E(C,\Sigma) = C \diagdown \Sigma^{-1}(C)$ and $I(C,\Sigma) = \Sigma(C) \cap \Sigma^{-1}(C)$; and, for each $\sigma = (x,y)$ in $C \times C$, we set $\sigma_- = x$ and $\sigma_+ = y$. Now, such a couple (C,Σ) will be called a *neuronal graph* if the following axioms are satisfied: $(NG1)A(C,\Sigma) \neq \emptyset$; $(NG2)A(C,\Sigma) \cap E(C,\Sigma) = \emptyset$; $(NG3)$ $\forall \, x \in C, (x,x) \notin \Sigma$. We note that $(NG2)$ is equivalent to $C = \Sigma(C) \cup \Sigma^{-1}(C)$, and imply that $\{A(C,\Sigma), E(C,\Sigma), I(C,\Sigma)\}$ is a partition of C. A neuronal graph (C',Σ') is called a *neuronal subgraph* of (C,Σ) if $C' \subset C \, \& \, \Sigma' \subset \Sigma$. Let us set then $W = \{N,L,S,D\}$, $W_* = \{L,S,D\}$, $W_0 = \{N,D\}$, $W_1 = \{L,S\}$, and $N = \{0,1,2,\ldots\}$; and let $\Omega(C,\Sigma)$ be the set of all $\omega \in Mp(N \times \Sigma, W)$‡ such that: (NE_1) $\forall \, \sigma \in \Sigma, \omega(0,\sigma) = N$; (NE_2) $\forall \, t \in N, \forall \, t' \in N$, $\forall \, \sigma \in \Sigma, t' \leq t \, \& \, \omega(t,\sigma) = N \Longrightarrow \omega(t',\sigma) = N$.

Neuronal graphs (defined above as mathematical structures) will now receive several different, although connected, interpretations (Int)† in terms of neuronal networks. *First*, a neuronal graph (C,Σ) provides a model of the static connective organization of a definite neuronal network: (Int 1) the synapses are labeled by the elements of Σ; (Int 2) the

---

† A model is considered as an axiomatically defined mathematical structure that is coupled with phenomenological premises and postulates about mechanisms by means of interpretative statements. The mathematical developments are carefully distinguished from their interpretations; the biological terminology used in those developments should not induce any modification whatsoever of the axiomatic setting of the models.

‡ $Mp(X,Y)$ denotes the set of all mappings from the set X into the set Y.

neurons are in one-to-one correspondence with the elements of $I(C,\Sigma)$; (Int 3) the soma, the axon hillock§, and the axon of the neuron $\nu(x)$ associated with $x \in I(C,\Sigma)$ are, respectively (resp), represented by $(\Sigma^{-1}(x),x)$, x, and $(x,\Sigma(x))$; (Int 4) the elements of $A(C,\Sigma)$ (resp of $E(C,\Sigma)$) represent the entry (resp exit)-hillocks; (Int 5) the elements of W_* represent the three possible connective states of the synapses (postulate **P2**), in such a way that each $\epsilon \in Mp(\Sigma,W_*)$ represents a connective organization of the network, with $\epsilon(\sigma)$ standing for the connective state of the synapses $\sigma \in \Sigma$ in that organization. More elaborate models of series of networks, somas with large dendritic arborizations (the Purkinje cell for instance) or special synapses (see *section 4.3*, Int 28), can be made in particular by modifying Int 1 to Int 3 and representing certain somas and synapses by neuronal subgraphs.

1.2 Evolution of Neuronal Networks. Second, a neuronal graph (C,Σ) can constitute a model of the geometry of the genetic envelope of a neuronal network (postulate **P4**) through the following interpretations: (Int 6) the genetically programmed maximum wiring is represented by (C,Σ) in such a way that the definite networks that can be actualized are in one-to-one correspondence with the elements of $Mp(\Sigma,W)$; (Int 7) each element of $A(C,\Sigma)$ (resp $E(C,\Sigma)$) represents a possible entry (resp exit)-hillock of the actualizable networks; (Int 8) the time scale (end of postulate **P1**) is represented by the ordered set N; (Int 9) each element ω of $\Omega(C,\Sigma)$ represents a complete evolution of the network including growth from time $t = 0$, with N for "Néant."

1.3 Neuronal Programs and the Basic Theorem of Neuronal Dynamics. Let $V = \{0,1\}$, $N^* = N \diagdown \{0\}$, and, for each $1 \in N^*$, let $]l[= \{1,2,\dots,l\}$ and $[l] = \{0,1,\dots,l\}$. Consider a mathematical structure $R = (C,\Sigma,\theta,n,\Phi,\Delta)$ where (C,Σ) is a neuronal graph, $\theta \in Mp(\Sigma,N)$, $n \in N^*$, Φ is a family $(\Phi_x)(x \in \Sigma(C))$ such that, for each $x \in \Sigma(C)$, Φ_x is a mapping from $Mp(]n] \times \Sigma^{-1}(x), V)$ into V, and Δ is a family $(\Delta_\sigma)(\sigma \in \Sigma)$ such that, for each $\sigma \in \Sigma$, Δ_σ is a mapping from the product set $N \times Mp(]n] \times \Sigma^{-1}(\sigma_+),W) \times Mp(]n] \times \Sigma^{-1}(\sigma_+),V)$ into W. Such a structure will be called a *neuronal program* if the following axioms are satisfied: $(NP1)$ $\forall\, x \in \Sigma(C)$, $\Phi_x(0_x) = 0$ (with $\forall\, j \in]n]$, $\forall y \in \Sigma^{-1}(x)$, $0_x(j,y) = 0$); $(NP2)$ $\forall\, \sigma \in \Sigma$, $\forall\, \xi \in Mp(]n] \times \Sigma^{-1}(\sigma_+),W)$, $\forall\, \mu \in Mp(]n] \times \Sigma^{-1}(\sigma_+),V)$, $\forall\, t \in N$, $\Delta_\sigma(t;\xi,\mu) = N \Rightarrow \xi(1,\sigma_-) = N$. Then, for each $t \in N$, $x \in \Sigma(C)$, $\omega \in \Omega(C,\Sigma)$, $U \in Mp(N \times C,V)$ and $U_- \in Mp(N \times \Sigma,V)$, let us define $\underline{\omega}[t,x] \in Mp(]n] \times \Sigma^{-1}(x),W)$ and $\underline{U}^\omega[t,x] \in Mp(]n] \times \Sigma^{-1}(x),V)$ by the following equalities (where $j \in]n]$ and $y \in \Sigma^{-1}(x)$):
$\underline{\omega}[t,x](j,y) = \omega(t + 1 - j,(y,x))$ if $t + 1 - j \geq 0$ and
$\underline{\omega}[t,x](j,y) = N$ otherwise; and
$\underline{U}^\omega[t,x](j,y) = U_-(t + 1 - j,(y,x))$ if $t + 1 - j \geq 0$
$\&\ \omega(t + 1 - j,(y,x)) \in W_1$,
and $\underline{U}^\omega[t,x](j,y) = 0$ otherwise.

With these notations, we shall say that (ω,U,U_-) is an *actualization* of the neuronal program R when the following equations are satisfied:

$(ND1)$ For all $t \in N$ and $\sigma \in \Sigma$, $U_-(t,\sigma) = U(t - \theta(\sigma),\sigma_-)$ if $t \geq \theta(\sigma)$, and $U_-(t,\sigma) = 0$ if $t < \theta(\sigma)$ or $t = 0$.
$(ND2)$ For all $t \in N$ and $x \in \Sigma(C)$, $U(t,x) = \Phi_x(\underline{U}^\omega[t,x])$;
$(ND3)$ For all $t \in N$ and $\sigma \in \Sigma$,
$$\omega(t + 1,\sigma) = \Delta_\sigma(t;\underline{\omega}[t,\sigma_+],\underline{U}^\omega[t,\sigma_+]).$$

§ The point of emergence of the axon from the soma.

By Eq. $(ND1)$, the actualization (ω,U,U_-) is determined by (ω,U): ω is called the *evolution* and U the *functioning*.

Now, the fundamental result of neuronal dynamics reads:
THEOREM: *Let* $R = (C,\Sigma,\theta,n,\Phi,\Delta)$ *be a neuronal program and let* $A = A(C,\Sigma)$. *For each* $\alpha \in Mp(N \times A,V)$, *there exists an unique actualization* (ω,U,U_-) *of* R *such that,*
(1.2) $\forall\, t \in N$, $\forall\, x \in A$, $U(t,x) = \alpha(t,x)$.
And, for each $t_0 \in N$, *the restriction of* U *(resp* ω) *to* $[t_0] \times C$ *(resp to* $[t_0] \times \Sigma$) *depends only upon that of* α *to* $[t_0] \times A$.

The proof is by induction on t along the graph (C,Σ). Note that the hypotheses imply that $\forall\, x \in \Sigma(C)$, $U(0,x) = 0$,

1.4 Actualizations of Neuronal Programs and Epigenesis of Neuronal Networks. A neuronal program R and its actualization (ω,U,U_-) can receive the following interpretations in terms of the whole genetic envelope (postulate **P4**) of an evolving neuronal network and of the resulting evolution by functioning: Int 6–Int 9 in regard to (C,Σ) and ω; (Int 10) functioning of the network (postulate **P1**) is represented by the binary wave (U,U_-) which stipulates the "informational value" (0 or 1) of the impulse at any given time and in any relevant place of the network (i.e., hillocks $x \in C$ and the afferent side of the synaptic cleft $\sigma \in \Sigma$); (Int 11) in particular, the restriction of U to $N \times A$ (resp $N \times E$) represents the total flow of impulses afferent to (resp efferent from) the network; (Int 12) for each $t \in N$ and $x \in \Sigma(C)$, $\underline{\omega}[t,x]$ (resp $\underline{U}^\omega[t,x]$) represents the evolution of the connective state of the synapses (resp the actual multimessage) afferent to the soma $(\Sigma^{-1}(x),x)$ during the time interval $\{j \in N | t + 1 - n \leq j \leq t\}$ (thus n appears as a somatic degree of memorization); (Int 13) for each synapse, $\sigma \in \Sigma$, $\theta(\sigma)$ represents (Eq. $ND1$) the delay of propagation of the impulse on the axon between the preceding hillock σ_- and the presynaptic terminal; (Int 14) for each $x \in \Sigma(C)$, the mapping Φ_x represents (Eq. $ND2$) the integrative power of the soma $(\Sigma^{-1}(x),x)$ (postulate **P4**); (Int 15) for each $\sigma \in \Sigma$, Δ_σ represents (Eq. $ND3$) the evolutive power upon the synapse σ of the soma $(\Sigma^{-1}(\sigma_+),\sigma_+)$ (postulates **P3** and **P4**) with emphasis on its purely local action. Suitable functions Φ_x could represent the firing mechanism and the excitatory or inhibitory character of the synapses; suitable functions Δ_σ could represent special constraints on the synaptic state transitions (see postulate **P2** and *section 4.2 d*), whereas suitable time dependance in Δ_σ could represent time unhomogeneous features of genetically programmed growth (postulate **P4**)

2. Neuronal functioning and behavior of biological organisms

2.1 Formal Organisms. Let R_+ be the set of non-negative real numbers. If $\mathfrak{X}$ is a topological space and $1 \in N^*$, let $T(\mathfrak{X})$ (resp $T_1(\mathfrak{X})$) be the topological space of all continuous mapping from R_+ (resp[0,1]) into $\mathfrak{X}$ with the compact convergence topology, and, for each $X \in T(\mathfrak{X})$ and $t \in N$, let $X_{[t]} \in T_1(\mathfrak{X})$ be defined by $\forall\, \tau \in [0,1], X_{[t]}(\tau) = X(\tau + t)$.

Now, a *formal organism* is defined as a structure $G = (\mathcal{S},\mathcal{Z},\mathfrak{X},s_0,\Gamma)$ where $\mathcal{S}$, $\mathcal{Z}$, and $\mathfrak{X}$ are topological spaces, $s_0 \in \mathcal{S}$, and Γ is a continuous mapping from $\mathcal{S} \times T_1(\mathcal{Z}) \times T_1(\mathfrak{X})$ into $T_1(\mathcal{S})$. It is then easy to prove that, given $Z \in T(\mathcal{Z})$ and $K \in T(\mathfrak{X})$, there exists a unique $S \in T(\mathcal{S})$ such that,
(2.1) $S(0) = s_0$ $\&$ $\forall\, t \in N$, $S_{[t]} = \Gamma(S(t), Z_{[t]},K_{[t]})$.

According to control theory, G might be interpreted in terms of concrete biological systems (muscle fiber, limb, entire body, etc): (Int 16) elements of $\mathcal{S}$ (resp of $\mathcal{Z}$) are in one-to-one

correspondence with the possible states of the formal organism (resp of its relevant environment); (Int 17) elements of $\mathcal{K}$ are in one-to-one correspondence with possible elementary *commands* given to the formal organism by its *control apparatus* (here a neuronal one; see *sections 2.2* and *2.3*); (Int 18) together with the dynamic equation (2.1), Γ represents the mechanism according to which the evolution S of the system results from the initial state s_0 and the evolutions Z and K of the environment and of the command in the physical time continuum R_+, all scanned by the time scale N (which will be that of the neuronal control apparatus; see Int 8, *sections 2.2* and *2.3*). Usual specifications of S, Z, $\mathcal{K}$ are manifolds, in general of infinite dimension (i.e., their elements are fields), whereas Eq. 2.1 results from evolution differential equations (see *section 4.2*, Eqs. **4.1** and **4.2**).

2.2 Dynamics of Neuronal Control Systems. Let R be a neuronal program (see *section 1.4*) such that $E = E(C,\Sigma) \neq \phi$, and let G be a formal organism (see *section 2.1*). A *coupling* of R with G is defined as a triplet (A',ψ,π), where A' is a subset of $A = A(C,\Sigma)$, ψ a mapping from $T_1(S) \times T_1(Z)$ into $Mp(A'',V)$ (with $A'' = A \smallsetminus A'$), and π a mapping from $Mp(E,V) \times \mathcal{K}$ into $T_1(\mathcal{K})$ such that,
(2.2) $\forall \xi \in Mp(E,V), \forall k \in \mathcal{K}, \pi(\xi,k;0) = k$.

The structure $Q = (R,G;A',\psi,\pi)$ will be called a *neuronal control system*, and a *closed* one if $A' = \phi$.

If $U \in Mp(N \times C,V)$, $S \in T(S)$, $Z \in T(Z)$ and $K \in T(\mathcal{K})$, we shall say that (U,S,K) is a *total actualization* of the neuronal control system Q with *outside evolution* Z when the following conditions are satisfied: (NB1) U is a functioning of R (see *section 1.4*); (NB2)$S(0) = s_0$ & $\forall t \in N$, $S_{[t]} = \Gamma(S(t), Z_{[t]}, K_{[t]})$; (NB3)$U_{A''}(0) = 0$ & $\forall t \in N^*$, $U_{A''}(t) = \psi(S_{[t-1]}, Z_{[t-1]})$; (NB4) $\forall t \in N, K_{[t]} = \pi(U_E(t), K(t))$, where for each $t \in N$ and $X \subset C$, $U_X(t) \in Mp(X,V)$ denotes the restriction of $U(t,.)$ to X.
Then, the fundamental result of the dynamic of neuronal control systems reads:
THEOREM: *Let* $Q = (R,G;A',\psi,\pi)$ *be a neuronal control system, and let* $\alpha' \in Mp(N \times A',V)$ *and* $Z \in T(Z)$. *Then, there exists a unique total actualization* (U,S,K) *of Q with outside evolution Z such that,*
(2.3) $\forall t \in N, \forall x \in A', U(t,x) = \alpha'(t,x)$. *And, for each* $t_0 \in N$, *the restriction of S to $[0,t_0]$ (resp of K to $[0,i_0]$, of U to $[t_0] \times C$) depends only upon those of Z to $[0,t_0]$ and of α' to $[t_0] \times A'$.*

Because of the fundamental theorem of neuronal dynamics (see *section 1.3*), the proof is elementary by induction on t.

2.3 Neuronal Control of Biological Systems. A neuronal control system $Q = (R,G;A',\psi,\pi)$, together with its total actualization (U,S,K) and its outside evolution Z, is intended to constitute a model of the *entire system* formed by an evolving neuronal network coupled with a biological system together with their evolutions by effect of the environment: Int 1–Int 15 (see *section 1*) in regard to R and U, and Int 16–Int 18 (see *section 2.1*) in regard to G, S, K, and Z; (Int 19) the restriction $\alpha' = U_{A'}$ of U to A' represents the flow afferent to the neuronal network from "the rest" of the nervous system; (Int 20) the mapping ψ represents the sense organs of the system that codes interoceptive (i.e., about S) and exteroceptive (i.e., about Z) sensory signals giving the afferent flow $U_{A''}$ (conditions $NB4$); (Int 21) the mapping π represents

the decoding by the terminal nerves of the efferent flow U_E from the neuronal network that gives a command for the coupled formal organism (condition $NB4$). Note that S can be large enough to let S include some structural evolution under the nonevolutive dynamic Γ, that S could influence Z (see *section 3.3 e*) and that the ψ and π mechanisms have no memory by themselves (every relevant memory being *neuronal*).

3. General features of neuronal learning

3.1 Three Fundamental Mappings. Let Q be a neuronal control system (see *section 2.2*) and $l \in N^*$. First, for each $\lambda \in Mp([l - 1] \times A,V)$ we define two mappings $\underline{R}_{[\lambda]}$ and $\underline{D}_{[\lambda]}$, with $\underline{R}_{[\lambda]} \in Mp(Mp(N \times A,V), Mp(N \times E,V))$ and $\underline{D}_{[\lambda]} \in Mp(Mp(N \times A,V), Mp(N \times \Sigma,W))$ by setting, for $\alpha \in Mp(N \times A,V)$, $t \in N$, $x \in E$ and $\sigma \in \Sigma$,
$\underline{R}_{[\lambda]}(\alpha;t,x) = U(t,x)$ and $\underline{D}_{[\lambda]}(\alpha;t,\sigma) = \omega(t,\sigma)$,
where $(\omega,U,U\text{-})$ is the unique actualization of R (see *section 1.3*) such that, for all $y \in A$, $U(t,y) = \lambda(t,y)$ if $t \in [l - 1]$ and $U(t,y) = \alpha(t - l,y)$ if $t \in N \smallsetminus [l - 1]$. Second, for each $z \in Z$, let $\varepsilon_{[z]}$ be the set of $(\alpha',Z) \in Mp(N \times A',V) \times T(Z)$ such that $Z(0) = z$; and let $\mathcal{L}_{[l,z]}$ be the set of $\Lambda = (\mu,M) \in Mp([l - 1] \times A',V) \times T_1(Z)$ such that $M(l) = z$. Then, for each $\Lambda = (\mu,M) \in \mathcal{L}_{[l,z]}$, we define the mappings $\hat{\Lambda} \in Mp([l - 1] \times A,V)$, $\underline{U}_{[\Lambda]} \in Mp(\varepsilon_{[z]}, Mp(N \times A,V))$, and $\underline{Q}_{[\Lambda]} \in Mp(\varepsilon_{[z]}, T(S))$ by setting, for each $(\alpha',Z) \in \varepsilon_{[z]}$, $t' \in [l - 1]$, $t \in N$, $x \in A$ and $\tau \in R_+$,
$\hat{\Lambda}(t',x) = U(t',x)$, $\underline{U}_{[\Lambda]}(\alpha',Z;t,x) = U(t + l,x)$, and $\underline{Q}_{[\Lambda]}(\alpha',Z;\tau) = S(\tau + l)$, where (U,S,K) denotes the unique total actualization of Q (see *section 2.2*) such that, for all $x \in A'$, $U(t,x) = \mu(t,x)$ if $t \in [l - 1]$ and $U(t,x) = \alpha'(t - l,x)$ if $t \in N \smallsetminus [l - 1]$, and with outside evolution $Z_{[M]}$ determined by $Z_{[M]}(\tau) = M(\tau)$ if $\tau \leq l$ and $Z_{[M]}(\tau) = Z(\tau - l)$ if $t \geq l$.

These definitions are interpreted in terms of neuronal learning for the entire system represented by Q (see *section 2.3*): (Int 22) the total input to the system during the learning procedure of duration l (the "*learning input*") is represented by an element $\Lambda = (\mu,M)$ of $\mathcal{L}_{[l,z]}$, the direct neuronal input (see Int 19) being represented by μ and the behavioral one (see Int 16) by M, with z for the state of environment at the end of the learning period (i.e., at time $\tau = l$); (Int 23) the total flow to the neuronal network (see Int 11) which results from the learning input Λ (see Int 22) is represented by $\hat{\Lambda}$; (Int 24) the ability of the connective organization to evolve and the functioning ability acquired by the neuronal network after the learning input Λ are represented, respectively, by the mappings $\underline{D}_{[\hat{\Lambda}]}\circ\underline{U}_{[\Lambda]}$ and $\underline{R}_{[\hat{\Lambda}]}$; (Int 25) the behavioral ability acquired by the entire system after the learning input Λ is represented by the mapping $\underline{Q}_{[\Lambda]}$.
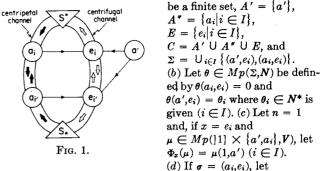
3.2 Effects of Environment. Thus, the effects of a learning input on the connective organization, on the functioning, and on the behavior, are, respectively, represented by the mappings: $\Lambda \to \underline{D}_{[\hat{\Lambda}]}\circ\underline{U}_{[\Lambda]}$, $\Lambda \to \underline{R}_{[\Lambda]}$, and $\Lambda \to \underline{Q}_{[\Lambda]}$. Now, it can be proved that, for suitable but not exceptional systems Q, none of these three quantities is a function of the two other ones, in particular, different learning inputs may produce *different* connective organizations and neuronal functioning abilities, but the *same* behavioral ability. Note that this *variability* at the functioning level with respect of the behavioral one appears in spite of the totally deterministic character of the model.

3.3 Associative Property: Neuronal and Behavioral Competences. Let us call *neuronal competence* of R [resp *behavioral competence* of Q] any subset $\mathfrak{N}$ of the set $Mp(Mp(N \times A, V), Mp(N \times E,V))$ [resp $z \in Z$ being given, any subset $\mathfrak{B}$ of $Mp(\mathcal{E}_{[z]}, T(\mathcal{S}))$; see *section 3.1*] of which elements have the causality property satisfied by every $\underline{R}_{[\lambda]}$ [resp $\underline{Q}_{[\lambda]}$] by virtue of the last statement of theorem *1.3* [resp *Theorem 2.2*].

Now, considering that an associative property can be represented in extended form by a set of associative abilities, we set: (Int 26) any associative property to be achieved by a learning process is represented by a neuronal competence $\mathfrak{N}$ of R at the functioning level, and by a behavioral competence $\mathfrak{B}$ of Q at the behavioral one, in such a way that (Int 23–Int 25) the actual achievement of the associative property $(\mathfrak{N},\mathfrak{B})$ after the learning input Λ is expressed by the relation $R_{[\hat{\lambda}]} \in \mathfrak{N}$ at the functioning level and by the relation $\underline{Q}_{[\Lambda]} \in \mathfrak{B}$ at the behavioral one.

Concerning the specification of competences, we make the following remarks: (a) A special class of neuronal competences (the "5-5'-competences") results when $\mathfrak{N}$ is defined as the set of $F \in Mp(Mp(N \times A,V), Mp(N \times E,V))$ having the causality property and such that $F(\mathfrak{I}_i) \subset \mathfrak{I}'_i$ for $i = 1, 2,\ldots, p$, where, for each i, $\mathfrak{I}_i$ (resp $\mathfrak{I}'_i$) is a given recursive subset (possibly defined in frequencies terms) of $Mp(N \times A,V)$ [resp $Mp(N \times E,V)$] which represents a type of afferent (resp efferent) message. In connection with the basic variability (see *section 3.2*), this class is probably too narrow to include all neuronal competences. (b) *Behavioral competences* (in particular homeostatic ones) should usually be defined in terms of a continuous causal mapping η from $T(\mathcal{S})$ into $T(\mathcal{K})$ where $\mathcal{K}$ is a given metric space. (c) Properties of periodicity (i.e., of time stability) have to be further introduced for competences, owing to their infinite extension in time (by definition) and of finiteness of the neuronic programs. (d) In opposition to the classical theory of control, the achievement of a competence has nothing to do here with any optimization principle whatsoever; it is only the expression of an epigenetic interaction. (e) Although the model in question does not fit with the stimulus–response theory mentality, it can nevertheless include reinforcement learning procedures by introducing a suitable influence of S onto Z. Two of the most critical features of the theory are not made explicit here: the *selective* aspect of learning and the presence of a *critical learning period;* both aspects can be formalized in the model and shall appear below in the application of the theory to the neuromuscular junction.

4. Application to the neuromuscular junction

4.1 Biological Premises. (a) At the stage of development of the neuromuscular junction referred to as "exploratory fibers," the contacts established between motor nerve endings and striated (fast) muscle fibers in high vertebrates are multiple and irregularly spread on the whole surface of the muscle fiber. (b) In the adult, a single endplate, in general, persists in the middle of the fiber. (c) Maturation, i.e., in particular the progressive degeneration of lateral contacts, is concomitant with nerve and muscle functioning. (d) Since significant fluctuations occur during embryonic nerve growth, the precise position of the endplate is not expected to be genetically programmed.

centripetal channel / centrifugal channel

FIG. 1.

4.2 Specification of a Neuronal Control System. (a) Let I be a finite set, $A' = \{a'\}$, $A'' = \{a_i | i \in I\}$, $E = \{e_i | i \in I\}$, $C = A' \cup A'' \cup E$, and $\Sigma = \cup_{i \in I} \{(a',e_i),(a_i,e_i)\}$.
(b) Let $\theta \in Mp(\Sigma,N)$ be defined by $\theta(a_i,e_i) = 0$ and $\theta(a',e_i) = \theta_i$ where $\theta_i \in N^*$ is given $(i \in I)$. (c) Let $n = 1$ and, if $x = e_i$ and $\mu \in Mp(]1] \times \{a',a_i\},V)$, let $\Phi_x(\mu) = \mu(1,a')$ $(i \in I)$.
(d) If $\sigma = (a_i,e_i)$, let
$\Delta_\sigma(.,.,.) \equiv S;$ if $\sigma = (a',e_i)$, $\zeta \in Mp(]1] \times \{a',a_i\},W)$ and $\mu \in Mp(]1] \times \{a',a_i\},V)$, let $\Delta_\sigma(0,\zeta,\mu) = L$ for all ζ and μ, and let $\Delta_\sigma(t,\zeta,\mu)$ be given by the following table according to the values of t, ζ, and μ:

	$1 \leq t \leq l_0 - 1$		$t \geq l_0$
$\zeta(1,a') = L$	$\Delta_\sigma(t,\zeta,\mu) = S$ if $\mu(1,a_i) = 1$		$\Delta_\sigma(t,\zeta,\mu) = D$ for all μ
	$\Delta_\sigma(t,\zeta,\mu) = L$ if $\mu(1,a_i) = 0$		
$\zeta(1,a') = S$	$\Delta_\sigma(t,\zeta,\mu) = S$ for all μ		
$\zeta(1,a') = D$	$\Delta_\sigma(t,\zeta,\mu) = D$ for all μ		

where $l_0 \in N$, $l_0 \geq 2$, is given. (e) Let $Z = \{0\}$, $S = C^\infty(R) \times C^\infty(R) \times C^\infty(R) \times C^\infty(R)$, $s_0 = (0,0,0,0)$, $\mathcal{K} = C^\infty(R)$ and let $\Gamma \in Mp(S \times T_1(\mathcal{K}), T_1(\mathcal{S}))$ (see *section 2.1*) be defined by the Cauchy problem for the coupled wave equations
(4.1) $v^{-2}\partial_0^2\varphi(\tau,\xi) - \partial_1^2\varphi(\tau,\xi) = K(\tau,\xi)$
(4.2) $v'^{-2}\partial_0^2\chi(\tau,\xi') - \partial_1^2\chi(\tau,\xi') = \varphi(\tau,\xi_*)H(\xi' - \xi_*) + \varphi(\tau,\xi^*)H(\xi' - \xi^*),$
$(\tau \in R_+, \xi \in R, \xi' \in R;$ with $\partial_0 = \partial/\partial\tau$ and $\partial_1 = \partial/\partial\xi)$ by setting
(4.3) $S(\tau) = (\varphi(\tau,.),\partial_0\varphi(\tau,.), \chi(\tau,.), \partial_0\chi(\tau..))$, where
$v > 0, v' > 0; \xi_* \in R, \xi_* \in R$ and $H \in C^\infty(R)$ are given such that
(4.4) $\xi_* < \xi^*$, $H \geq 0$, $H(-\xi) = H(\xi)$, H vanishes outside $[-\rho,\rho]$ (with $\rho > 0$ given) and $H(0) > 0$.
(f) Let $(\xi_i)_{i \in I}$ be a family of elements of $]\xi_*,\xi^*[$ such that $\xi_i \neq \xi_{i'}$ if $i \neq i'$, and let $\psi \in Mp(T_1(\mathcal{S}),Mp(A'',V))$ be determined, considering Eq. 4.3 and $\rho' > 0$, $\gamma > 0$ being given, by
(4.5) $\forall i \in I$, $\forall S \in T_1(\mathcal{S})$, .
$\psi(S;a_i) = 1 \iff \exists \tau \in [0,1], \exists \xi \in [\xi_i - \rho',\xi_i + \rho'], \chi(\tau,\xi) \geq \gamma$
(g) Let $\pi \in Mp(Mp(E,V) \times \mathcal{K}, T_1(\mathcal{K}))$ be determined by
(4.6) $\forall \delta \in Mp(E,V)$, $\forall k \in \mathcal{K}$, $\forall \tau \in [0,1]$, $\forall \xi \in R$,
$\pi(\delta,k;\tau,\xi) = k(\xi) + \Sigma_{i \in I} \delta(e_i)J(\tau)H(\xi - \xi_i)$,
where $J \in C^\infty([0,1])$ is given such that (with $0 < \epsilon' < \epsilon < 1$),
(4.7) $\int_0^1 J(\tau)d\tau = 0$ and $\forall \tau \in [0,\epsilon'] \cup [\epsilon,1], J(\tau) = 0$;

4.3 Interpretations. The neuronal control system Q described in *section 4.2* constitutes a model of the entire system (see *section 2.3*) made up of a motoneuron coupled in an evolutive manner with a muscle fiber (see *section 4.1*). Considering an isometric fiber contraction, we postulate that signals are transmitted through two distinct channels: action potentials initiated at the endplates are propagated through a "centrifugal" channel; whereas a "centripetal" channel transmits, back to the endplates, the signals initiated at the ends of the fiber upon arrival of the action potential (see *section 4.2* and Fig.); (Int 27) the axon (resp the axon hillock) of the motoneuron is represented by (a',E) (resp by a'); its soma is not represented; (Int 28) the endplates (see *sections 4.1 a* and *b*) are labeled by the elements of I; the *i*th endplate is represented by the neuronal subgraph $(\{a_i,e_i\}, (a_i,e_i))$ of (C,Σ) (see the end of *section 1.1*) with ξ_i for its abscissa along the axis of the

fiber; (Int 29) the formal organism $G = (\mathcal{S}, \mathcal{Z}, \mathcal{K}, s_0, \Gamma)$ represents the two channels considered to be unidimensional: the propagation of the action potential φ in the centrifugal channel is represented by the wave equation (4.1) (with propagation speed $v^{\P}$), that of the (still unidentified) signal χ in the centripetal channel by Eq. 4.2 (with propagation speed v'); the second member of Eq. 4.2 represents the initiation mechanism of the centripetal signal χ when the action potential φ arrives at the ends of the fiber (of abscissal ξ^* and ξ_*) (note that no reflection occurs at the ends of both channels); (Int 30) the mapping π together with the setting of K in the second member of Eq. 4.1 represents the mechanism by which the impulse coming from the motoneuron generates an action potential on the muscle fiber by the endplates; (Int 31) the mapping ψ represents a threshold mechanism (see Eq. 4.5) by which each endplate decodes the centripetal signal: the threshold γ (which is the same for all $i \in I$) is reached only when two centripetal signals cross near ξ_i; and the synaptic contact (a', e_i) is stabilized when such an event occurs during the critical period of synaptic plasticity represented by $[l_0 - 1]$ (see *section 4.2 d*).

4.4 Neuromuscular Junction Learning. For each $p \in \,]l_0 - 1]$, let $\mathfrak{I}(l_0, p)$ be the set of $\Lambda \in Mp([l_0 - 1] \times A'V)$ such that
(4.8) $\exists\ t \in [l_0 - 1 - p], \Lambda(t, a') = 1$, and
(4.9) $\forall\ t \in [l_0 - 1 - p], \forall\ j \in \,]p], \Lambda(t, a') = 1 \Longrightarrow \Lambda(t + j, a') = 0$;
and let $\Lambda_0 \in Mp([l_0 - 1] \times A', V)$ be the "null" learning input (i.e., $\forall\ t \in [l_0 - 1], \Lambda_0(t, a') = 0$). Furthermore, let us introduce the following properties: (H1) $v' < v$;
(H2) $\xi^* - \xi_* \geq (2v\epsilon + 4\rho + \rho') + v'\bar{\theta})/(1 - (v'/v))$, where $\bar{\theta} = \sup\{|\theta_{i'} - \theta_{i''}|\ |i' \in I, i'' \in I\}$; (H3) $l_0 \geq 4(\xi^* - \xi_*)/v'$;
(H4) $\forall\ i' \neq i'', |\xi_{i'} - \xi_{i''}| \geq 2\rho$;
(H5) $\forall\ i, |\xi_i - \xi_m| \leq \xi^* - \xi_m - (v'\epsilon + 4\rho)$, where $\xi_m = (\xi^* + \xi_*)/2$; (H6) $p \geq 3(\xi^* - \xi_*)/v'$. Then:

THEOREM. *Beside the general conditions stated in section 4.2, let us suppose that H1, H2, and H3 are satisfied. Then, there exist functions J and H satisfying Eqs. 4.4 and 4.7 and $\gamma > 0$ such that, for every $(\xi_i)_{i \in I}$ satisfying H4 and H5, and every $p \in [l_0 - 1]$ satisfying H6,*
(A) for all $\Lambda \in \mathfrak{I}(l_0, p)$, $\alpha' \in Mp(N \times A', V)$ and $t \in N$,
(4.10) $\omega(\Lambda, \alpha'; t, \sigma_i) = S$ for all $i \in I$ such that
(4.11) $|\xi_i - \xi_m| \leq \rho'/(1 + (v'/v))$, and
(4.12) $\omega(\Lambda, \alpha'; t, \sigma_i) = D$ for all $i \in I$ such that
(4.13) $|\xi_i - \xi_m| \geq v'\epsilon + 4\rho + \rho' + v'\bar{\theta}/2 + (v'/v)(\xi^* - \xi_*)/2$,
where $\sigma_i(a', e_i)$, and $\omega(\Lambda, \alpha') \in Mp(N \times \Sigma, W)$ is defined by
(4.14) $\omega(\Lambda, \alpha') = D_{[\hat{\Lambda}]}(\underline{U}_{[\Lambda]}(\alpha'))$ (see section 3.1);
(B) for all $\alpha' \in Mp(N \times A', V)$ and $t \in N$,
(4.15) $\omega(\Lambda_0, \alpha'; t, \sigma_i) = D$ for all $i \in I$.

¶ We have not attempted to describe the propagation of action potential in terms of transport of ions.

From the explicit fundamental solution of the one-dimensional wave equation, J and H can be chosen such that, for each isolated incoming impulse (see Eqs. 4.8 and 4.9, and H6) the solutions φ and χ of Eqs. 4.1 and 4.2 are spike-shaped signals and χ is non-negative. Dealing with these signals, the existence of the threshold γ can be proven, and the $\underline{D}_{[\hat{\Lambda}]} \circ \underline{U}_{[\Lambda]}$ of the entire system Q (see *section 2.1*) can be approximated by the mapping $D_{[\Lambda]}$ of the "fictitious" neuronal program $\tilde{R}$ defined, extending R (see Fig.), by $\tilde{C} = C \cup \{s^*, s_*\}$;
$\tilde{\Sigma} = \Sigma \cup \cup_{i \in I}\{(e_i, s^*), (e_i, s_*), (s^*, a_i)(s_*, a_i)\}$ (note that $A(\tilde{C}, \tilde{\Sigma}) = A'$ and $E(\tilde{C}, \tilde{\Sigma} = \emptyset)$; if $x \in \{s^*, s_*\}$ and $\mu \in Mp(]1] \times E, V)$, $\tilde{\Phi}_x(\mu) = \sup\{\mu(1, e_i)|i \in I\}$; if $x = a_i$ $(i \in I)$, and $\mu \in Mp(]1] \times \{s^*, s_*\}, V)$, $\tilde{\Phi}_x(\mu) = 1$ if $\mu(1, s^*) = \mu(1, s_*)$; if $\sigma \in \tilde{\Sigma} \setminus \Sigma$, $\tilde{\Delta}_\sigma(., ., .) \equiv S$.

The preceding theorem shows that the theory accounts for the biological premises (see *section 4.1*) if $v'\epsilon$, ρ, ρ', and $v'\bar{\theta}$ are sufficiently smaller than the fiber length and if v'/v is small: for every input $\Lambda \in \mathfrak{I}(l_0, p)$, where p is large (see H6), only the endplates that are sufficiently close to the middle of the fiber are selectively stabilized (see Eqs. 4.10–4.15) and this phenomenon occurs whatever the abscissal $\xi_i (i \in I)$ of the endplates, as long as they are not too close to one another and to the ends (see *section 4.1 d*). On the contrary, in the absence of functioning during the critical period $[l_0 - 1]$, all the synaptic contacts degenerate (see *section 4.1 c*). Coming back to our initial distinction between the two levels at which learning should be considered, we emphasize that the achieved neuronal competence after a learning input $\Lambda \in \mathfrak{I}\,(l_0, p)$ is of the $\mathfrak{I}$-$\mathfrak{I}'$ type (see *section 3.3 a*), where as the behavioral one can be expressed as the symmetrical character of the muscle-fiber contraction $\eta(S)$ (suitably formalized according to *section 3.3 b*) after degeneration of the lateral contacts. Here, the learning process of the relevant system (motoneuron + muscle fiber) does not derive from an interaction with its environment but results strictly from the neuronal input (see *section 2.3*, Int 19).

1. These postulates are discussed in: Young, J. Z. (1964) *A Model of the Brain* (Oxford Univ. Press, London); Changeux J. P. (1970) *La Recherche* **1**, 213–221; Changeux, J. P. (1972) *Communications* **18**, 37–47; Changeux, J. P. & Danchin, A. (1972) in *The Limits of Man* (C.I.E.B.A.F., Royaumont, Sept. 1972).

August 23, 1973

Dr. Antoine Danchin
Institut Pasteur
25 rue du Dr. Roux
75015 Paris, France

Dear Dr. Danchin: Re: Manuscript #670

   Your manuscript has been estimated in the PROCEEDINGS Office to
run slightly over the 5-page limit. Rather than delay the manuscript at
this stage, we have edited the paper with an eye toward bringing the
length down to 5 pages. If there have been any distortions of your
meaning introduced in the process, please feel free to correct them at
the galley proof stage, bearing in mind the limitation of 5 pages.

                    Sincerely,

                    Murrie W. Burgan

                    Murrie W. Burgan, Ph. D.
                    Senior Associate Managing Editor

MWB:cw