Self-organization in the nervous system

J. Szentágothai and P. Erdi

1st Department of Anatomy, Semmelweis University Medical School, Budapest IX. Tuzolto u. 58, H-1450, Hungary and 2Central Research Institute for Physics, Hungarian Academy of Sciences, H-1525 Budapest, P.O.B. 49, Hungary

The usefulness—if not heuristic value—of the notion of 'self-organization' is examined in order to understand the 'essence of neural'. Self-organization is considered as a mechanism for generating emergent complex structures. One of its most striking properties is the capability to create order from noise. Self-organization phenomena are illustrated in very simple and highly complex neural structures. An attempt is made to interpret development, learning, and perhaps even dynamic memory organization in terms of neurodynamic principles. Ontogeny, normal performance, and plasticity of the nervous system are viewed in the framework of conceptually coherent phenomena and/or principles. The 'essence of neural' ought to be reconsidered before continuing the discussion of the brain–mind problem.

1. Introduction

Although vague in many respects, the idea of 'self-organization' is nevertheless a powerful concept of modern theoretical sciences. The concept of self-organization is associated with more or less abstract notions, e.g. macroscopic order and disorder, pattern formation (and recognition), dissipative structures, synergetic systems, emergence of complexity, coherence, order through fluctuations, adaptability, thermodynamic and information— theoretical entropy, noise—induced transition, autopoietic systems etc. (Atlan, 1981, 1985; Brillouin, 1962; Casti, 1986; Casti & Karlquist, 1986; Conrad, 1983; Glansdorff & Prigogine, 1971; Haken, 1985a, b; Horsthemke & Lefevre, 1984; Jantsch, 1980; Maturana & Varela, 1980; Nicolis & Prigogine, 1977; Rosen, 1981, 1986; Thom, 1972; von Foerster, 1960).

Self-organization, as a conceptual framework, proved to be efficient for explaining (sometimes implicitly) biological phenomena. Both ontogenetic development and phylogenetic evolution are dynamic processes to be identified with self-organization phenomena. Motivated by the abstract concept of a replicator, mathematical models for the prebiological evolution were extended to cover phenomena from other fields of biomathematics, such as population genetics, ecology and sociobiology. Even the behaviour of the immune system depends on the interaction between genetically determined mechanisms and environmental signals and noise (e.g. Dawkins, 1982; Eigen, 1971; Eigen & Schuster, 1979; Gould, 1977; Jerne, 1967; Schuster & Sigmund, 1983; Schuster, 1985; Sigmund, 1986).

The convergence between empirical data coming from neurobiology and approaches of dynamic system theory suggests that neural phenomena occurring at different
Hierarchical levels can be interpreted in terms of self-organization. Specifically, random spontaneous activity (on the macroscopic, i.e. supermolecular, level) might lead to temporal order (i.e. limit cycle type bursts of activity) by self-organization in the absence of any interference (let along selection) from/by the environment. These activity patterns may reach high levels of specificity, provide that the mutual connectivity in these self-contained neural networks is not random, but determined by some genetic algorithm. Its 'clockwork' is being driven essentially by internal mechanisms.

Our intention in this paper is to demonstrate that self-organization, viewed as a temporal process, whereby an essentially structureless system becomes 'organized', by virtue of nonlinear (partially random) dynamic processes, is a fundamental concept for a theoretical framework in the interpretation of neural phenomena. In Section 2 some theoretical concepts of the self-organizing systems are given. The relevance of the idea of self-organization in the nervous system is illustrated in Section 3 reviewing the basic experimental facts, while its theoretical power is demonstrated by treating ontogenetic development, anatomical and physiological plasticity, and psychological learning in a common framework (Section 4). Our final conclusion (Section 5) is that new aspects of the essence of neural organization should be clarified before continuing the discussion of the age-old brain–mind problem.

2. Self-organization: some concepts

Entropy and information

The term entropy, one of the vaguest notions of modern science, defined originally in the framework of phenomenological and statistical thermodynamics, has invaded many

![Derivation Diagram](image)

Fig. 1. Derivation (Érdi, 1983) of the relationship between the thermodynamic and information-theoretical entropy change. \( \pi \) denotes the thermodynamic entropy production, considered as a stochastic process. \( Q \) is the distribution of \( \pi \), \( E \) is the Kullback–Renyi entropy with negative sign. Further explanation in the text.
Self-organization in the nervous system

fields, as information theory, biology, behavioural science (Bateson, 1972), economy (Georgescu-Roegen, 1971) and socioeconomy (Wicken, 1986).

There is no agreement about whether entropy as an information-theoretical notion can be identified with thermodynamic entropy, and if so, in what sense. Brillouin's 'Negentropy Principle' that suggests the almost reversible interconvertibility of all kinds of information to thermodynamic entropy, has been questioned by Denbigh (Brunet, 1986; Brillouin, 1962; Denbigh, 1981; Kubat & Zemen, 1975).

A relationship between thermodynamic entropy and a precisely defined information-theoretical entropy was demonstrated earlier (Erdi, 1983) expressing the connection between all of the possible states and the actual state (Figure 1) adopting a macroscopic stochastic description.

It is necessary to realize that when neuroscientists speak of information processing in the nervous system, they generally use the term in a colloquial sense. Formal, Shannonian or syntactical information theory has a limited, but unquestionable use in neurophysiology. Tsuda & Shimizu (1985) gave a model for the mechanism of the transmission of both Shannonian and non-Shannonian information through neurons considered as dynamic channels. The nervous system is able not only to store and process, but to generate semantic information by self-organization (Shimizu, Yamaguchi, Tsuda & Yuno, 1985). The main difficulty with the concept of biological information comes from the self-referential character of biological systems. The relationship between system and observer has to be taken into account (Maturana & Varela, 1980; see also Kampis, 1987).

Complexity and redundancy

Thermodynamic entropy reflects the character of the molecular arrangement. Though other kinds of complexity might have similar significance, earlier attempts for finding a general expression based on the analogy of statistical physics, were not successful.

Keeping our eyes on the problem of complexity of the brain, at least three different notions of complexity may be defined. First, (structural) complexity can be associated to the number of cells and synapses to their ratio, furthermore qualitatively to the occurrence of many highly re-entrant loops, etc. Second (behavioural or functional) complexity can be identified with different dynamic behaviour determining the qualitative nature of the attractor of the process (point attractors, closed curves, strange attractors are mentioned, as complexity increases). Third (algorithmic) complexity can be measured with the minimal time needed to compute the solution of a problem. (The metaphor of classical cybernetics that the brain could be considered as a computer, although not quite fortunate, may have some practical relevance in its 'reversed form'. Based on our neurobiological knowledge about the organization of the brain, fault-tolerant, adaptive, parallel brain-like structures capable of 'intelligent' behaviour may be designed (Arbib, 1987; Caudill & Butler, 1987).

The problem of finding 'measures of complexity' has become interesting again in the light of the results on cellular autoata. The complexity of 'self-generated' spatiotemporal patterns (the space was also one-dimensional) has been examined (Grassberger, 1986a, b). Grassberger compares the entropy of three different patterns (Figure 2). The first is completely ordered, the second is generated by a simple cellular automat, the third is completely random. If complexity were associated to randomness, the third pattern would be the most complex. However, intuitively the second pattern could be qualified as the most complex, because it has some 'structure'. The contradiction has been resolved by Grassberger's suggestion of a new measure of complexity: '... complexity in the naive sense measures the information content of ensembles and not of individual patterns ...'. The theoretical physicist Grassberger's intuitive feeling about the mechanism of pattern recognition is in accordance with the 'resonance' and 'hermeneutics' hypotheses.
J. Szentágothai and P. Érdi

Fig. 2 Three patterns used to demonstrate that the pattern that one intuitively would call the most complex (b) is neither the one with highest entropy [and Kolmogorov complexity] (c) nor the one with lowest (a). After Grassberger (1986a).

A structure may be qualified as redundant when the arrangement of its elements shows repetitions. Repetition, though, is no equivalent to physical periodicity, in the simple sense of one element repeating itself many times. It can mean algorhythmic repetition: knowledge of one element gives some information on the others, and in certain sense this is the source of the emergence of order. Sometimes redundancy is referred to as the measure of simplicity. According to Shannon's definition the relationship between the information content $H$ and the redundancy $R$ is expressed as $H = H_{\text{max}}(1 - R)$, where $H_{\text{max}}$ is the maximal information content computed without taking into consideration the constraints between the parts of the system. Based on the equation referring to the temporal change of $H$, Atlan (1981, 1985) considers self-organizing system as 'a system redundant enough and functioning in such a way that it can sustain a decrease in redundancy under the error-producing factors, as fluctuations or noise, however, can be directly switched to self-organization, according to the theory of noise-induced transitions'.

**Noise induced transitions**

The formation of spatiotemporal structures may be neither the result of some rigid deterministic mechanism, nor the eventual outcome of a random series of events. According to general assumption, random fluctuations are 'averaged out', and macroscopic structures essentially follow deterministic laws.

The theory of statistical physics, shows that in certain situations, e.g. in the neighborhood of instability points, the effect of fluctuations cannot be averaged out, on the contrary, they may cause drastic effects on the macroscopic level. Fluctuations (noise) superimposed on deterministic laws can have a crucial role in forming ordered structures. According to the theory of noise-induced transitions (Horsthemke & Lefevre, 1984) fluctuations may operate as 'organizing forces'. In somewhat more technical terms, noise can destabilize the deterministic attractors, therefore stochastic models can exhibit properties that are qualitatively completely different from those of their deterministic counterpart.

The states of many macroscopic systems can be characterized by a finite-dimensional vector $C(t)$ at a fixed time $t$. The behaviour of such a system over time can be described by ordinary differential equations of the form

$$\frac{dC(t)}{dt} = f[C(t), t, \lambda].$$

The equation may include explicit time-dependence. The 'forcing' function $f$ also depends on a set of control parameters $\lambda$. In the most simple case, $\lambda$ is a scalar quantity. Whatever our definition of the vague concept of 'environment', it is plausible to
postulate that it has a random character. Leaving aside mathematical technicalities, a possible stochastic counterpart of the deterministic model mentioned above is

$$d\tilde{\mathcal{C}}_t = \left[ h(\tilde{\mathcal{C}}_t, t) + \lambda g(\tilde{\mathcal{C}}_t, t) \right] dt + \sigma g(\tilde{\mathcal{C}}_t, t) dW_t$$  \hspace{1cm} (2)$$

This is a stochastic differential equation, where the forcing function is the first two terms of the right-hand side, and \( W_t \) is the Wiener process.

The distribution function or probability density provides a practically sufficient characterization of a random variable, while the shape and modality (associated with the number of extrema) of the probability density function characterizes the system qualitatively. If the effect of the environmental fluctuation depends on the state of the system, the qualitative properties of the probability density function can be modified by increasing the environmental noise. Noise can increase the number of extrema, and in a certain sense leads to greater complexity.

The theory of noise-induced transitions has been applied conceptually for describing the ontogenetic development and plastic behaviour of ordered, neural structures, such as: retinotectal connections (Érdi & Barna, 1984), and ocular dominance columns (Barna & Érdi, 1986; Érdi & Barna, 1987).

Activity patterns heavily influence the formation and plastic behaviour of ocularity domains. It was demonstrated (Fraser, 1985) that treatment with tetrodotoxin of one eye of a kitten from 2 to 6 weeks of age prevents the normal development of ocular dominance columns. Since ocular dominance columns develop at least to some extent before birth in monkeys, if activity does play a role in columnar segregation, spontaneous activity (having random character) must be sufficient (Stryker, 1982; Stryker & Harris, 1986). A similar concept, the ‘order from noise’ principles introduced by von Foerster (1960) and modified by Atlan (1981) is based on the assumption that noise can reduce the redundancy of the system.

**Stability**

The formation and maintenance of dynamic structures is closely connected with the question of their stability. The issue of stability of structures has been studied from the point of view of topology (Thom, 1972), of thermodynamics (Glansdorff & Prigogine, 1971), and of mathematical system theory (Siljak, 1978). The criteria for the stability of the equilibrium states of deterministic, single-level systems were given by Liapunov a long time ago. These kinds of stability theories examine the effect of the change of the initial conditions and the parameter perturbations, but neglect the effect of ‘structural’ changes.

Biologically relevant dynamic systems are generally considered as heterogeneous assemblies of interacting constituents. During operation a group of the constituents can be disconnected from the rest of the system, causing a change in its morphology. These disconnections may be the results of faults in the natural behaviour of the system. In order to maintain proper functions, we should be able to answer questions of the following type: to what extent is the system’s performance affected by structural changes? Is it possible to secure stability of the system under preselected classes of structural perturbations? . . . (Siljak, 1978, p. 62). The notion of connective stability introduced by Siljak is appropriate for extending the notion of stability in Liapunov’s sense for accommodating structural perturbations. A connective stable system is dynamically reliable because it is stable under structural perturbation, whereby constituents may become disassociated and reassociated during the operation of the system.

In dealing with the nervous system, the stability–plasticity dilemma is particularly important (see Grossberg, 1988): ‘. . . How can a system’s adaptive mechanism be stable
enough to resist environmental fluctuations that do not alter its behavioural success, but
plastic enough to change rapidly in response to environmental demand that do alter its
behavioural success? How are stability without rigidity and adaptability without chaos
achieved? . . .' (p. 85).

3. Neural organization

Very simple neural systems

The intact nervous systems of animals, even in their simplest forms, are much too
complex to permit raising really fundamental questions. It is possible, though, to
assemble experimentally very simple systems in which the number of elements can be
reduced at will—at least at the system's level; (although not at the cellular level that
would need an entirely different approach). The simplest possible and also the most
elegant approach to such questions is tissue culturing (Crain 1973; Provine & Rogers,
1977) and, most important, the recent studies of Gähwiler (see for example Gähwiler &
Brown, 1987, and many other publications), in which hippocampal pyramidal neurons
cultured in tissue slices, or co-cultures of hippocampal and slices of other centres
(normally connected with the hippocampus) led this author to the conclusion that a
hitherto unexpected degree of *self-organization, adaptation, and plasticity* is present in
organoid tissue cultures. Since many of the elementary neuro-physiological mechanisms
have been clarified by these (and many other) studies, we need not worry about the basic
facts, but can turn back to earlier more phenomenological studies with tissue recombi-
nation experiments.

Using a technique originally devised by Paul A. Weiss (1950), small groups of nerve
cells can be replanted into virtually nerve-free regions of host animals [the dorsal fin of
Urodele larvae (*Triturus or Salamander*)] and a limb bud implanted nearby can be used
as an indicator of function. Our experiments from some 25 years ago (Szekely &
Szentagothai, 1962) and some later additional experiments are summarized in Figure 3.

Random groups of deplanted neurons of the early medullary tube [Figure 3 (a)] can
innervate the limb regenerated from the nearly implanted limb bud and elicit completely
random movements that have no resemblance whatever with natural movements. The
elementary conditions of such 'model nervous systems' are as follows:

1. Prospective motoneurons had to be present in the deplant, because only cholinergic
motoneurons can successfully innervate muscle fibres.

2. Other neurons (interneurons) had to be present (these conditions are *a priori* met,
because of the technique of excising tissue from the ventral quadrant of the early
tube), some of which were inhibitory interneurons. This was shown by the marked
strychnine effect (strychnine is an inhibitor of inhibition exercised upon moto-
neurons) seen in such models.

3. No sensory input was necessary, the activity was 'spontaneous'. (True reflex
models can be made by including sensory neurons from the spinal ganglia, but this
is irrelevant here and would unnecessarily complicate matters.)

4. Ten to twenty neurons randomly connected are sufficient; no organized structure
is necessary.

Natural-like movement can be produced by models (Szekely & Czeh, 1971) if the
following conditions are met [Figure 3 (b)]:

1. A certain degree of normal connectivity (wiring) had to be maintained. This was
secured by deplanting neural tube fragments keeping the central canal intact. As
known from modern neuroembryology, the inner lining of the neural tube
Fig. 3 Summary of tissue recombination experiments with isolated neural centres. Arrangement at left margin, showing the deplantation of neural centres and limb primordia into the dorsal fin of Triturus embryos. The drawing shows the situation in late larval stadia, when observations were made. (a) Experiment of Szekely & Szentágothai (1962) described in the text. The deplanted centre consists of a few (15–25) scattered nerve cells, the motoneurons, which innervate the muscles of the implanted limb. Although the cells are mixed in reality, the motoneurons (in stippled area) have been separated in the diagram because they served only as indicators. The generator of activity is shown at left as a randomly interconnected network of excitatory (outlines) and inhibitory interneurons (black). (b) Experimental arrangement of Szekely & Czeh (1971) where the deplanted centre is a segment of the medullary tube ventral horn. If taken from the limb-bearing segments of the medullary tube, the innervated limb graft exhibits the cyclic ‘walking-type’ movements of the limb illustrated by the circularly arranged heavy arrows; starting from the right: ‘lift limb’, ‘pull forward’, ‘press down’, and ‘push back’. (c) The ‘parasitic trunk-segment preparation of Brändle & Szekely (1973) illustrating the simple fact that the isolated limb segments of the medullary tube can move a pair of limbs only in parallel oarstroke-like manner; for the characteristic alternating movement pattern an additional part of the medullary tube, rostral from the limb segments, is necessary.
contains the cell material from which both the neurons arise and certain glial cells that serve as scaffolding for the migration of various types of neurons (Rakic, 1981). Consequently the neural deplant that can move a limb in quasi-step-like has to retain a central canal and a certain degree of organoid structure.

(2) The fragment had to be taken from segments of the medullary tube that would innervate limbs in the normal course of development. Non-limb-segments cannot generate 'stepping' type movements, as correctly inferred already from various earlier experiments).

(3) Sensory input was not needed in such models.

If double implants from both ventral halves of the medullary tube were implanted with two limb buds nearby, the limbs moved always synchronously in 'oar-stroke' fashion. In order to achieve the natural alternation of the limb pair, some additional (non-limb) segments of the medullary tube had to be present (Brändle & Szekely, 1973). Fig. 3(c) shows the somewhat different design of this experiment, which—unfortunately—does not show whether sensory input is necessary or not, but from circumstantial evidence it may be assumed that it is not.

**Conclusion**

Few randomly interconnected neurons (mainly interneurons) fully suffice to constitute a very simple neural system. No sensory input is needed in such elementary models for orderly function. Lacking sensory input the activity has, obviously, to originate from the known spontaneous activity of neurons. Neurons become spontaneously active, especially if isolated, but many neurons are known to be active spontaneously in their normal environment. A combination of excitatory and inhibitory neurons is probably needed, otherwise the system would turn immediately into a runaway excitatory feedback—as can be experimentally induced by strychnine, or seen under pathological circumstances in tetanus and in epileptic seizures. Also the models of Wilson & Cowan (1973) suggest the necessity of a combination of excitatory and inhibitory neurons. This condition is satisfied practically everywhere in neural subsystems [all segments (or any fragments) of the medullary tube, major subcortical nuclei, and in all cortical centres].

Since spontaneous activity of neurons is usually random, an organized activity can emerge only by self-organization of random neural noise under the experimental conditions summarized above. Given organoid (i.e. nature-like) connectivity of various types of neurons, highly specific output patterns—like stepping rhythm, or, with some additions, alternate stepping—can be produced in very elementary neural system models. Modern neurophysiological methods applying multiple microelectrode techniques can give information about the emergence of spatiotemporal coherence among the cells of neuron populations (Krueger, 1983, 1985).

The study of *global* brain activity dynamics by electroencephalography (EEG) gave evidence that spatiotemporal patterns of cortical neural activity are internally generated representations of expected sensory input and not merely responses to such input. Following this line (Freeman, 1983, Freeman & Skarda, 1985) the assumption can be made that the macroscopic cooperative activity of weakly but widely interconnected neurons emerges from the operations of individual neurons coupled by conventional synaptic transmission. Additionally, the construction of a perceptual representation from sensory input can be described in terms of self-organization.

There is no reason whatever to assume that these basic properties of simple artificial neural systems are lacking in the intact nervous system.

The observed autorhythmicity of embryonic neural tissues and its increase during embryogenesis was defined electrophysiologically by Llinas (1984) as cell-to-cell coupling by electrotonic synapses, assumed to be more frequent in the early embryonic neural
networks. The first step in the establishment of elementary networks would be assembly by random contacts and/or chemotaxis, this would lead in turn to electronic coupling of the intrinsic autorhythmicity of individual nerve cells into so-called ‘coupled oscillators’.

Admittedly our mental strategy includes the still somewhat obscure concept of ‘self-organization’, but it has the advantage of being more parsimonious by not having to rely on coupled oscillators. Such units cannot be considered as elementary building blocks, since they can exhibit complex dynamic behaviour. As is well-known, depending on the frequency ratio of the oscillators such phenomena as phase-locking (i.e. synchronicity), quasi-periodicity and even chaotic behaviour can appear.

The most complex neural structures

Experimental facts resulting from anatomy, physiology, embryology, and psychophysics give evidence in the vertebrate nervous system of highly ordered structure composed, as it were, of building blocks of repetitive structure. The building blocks, or ‘modular architectonics’, principle is quite common in the nervous system. The modular architecture of the spinal cord, of the brain stem reticular formation, of the hypothalamus, of the subcortical sensory relay nuclei, of the cerebellar and cerebral cortex has been reviewed recently by Szentagothai (1981, 1983, 1987).

The modular architectonics principle is essentially the result of the development of the vertebrate central nervous system. The medullary tube (or the brain vesicles as its rostral end), originally a single pseudostratified epithelium are giving rise to two major lineages of cells: neuroblast and glioblasts in the innermost layer of the tube (the so called ventricular zone), where all cell divisions take place. Since all cells have originally an attachment at the inner (ventricular) surface and another on the outer surface of the tube (lost for most cell types in later stages), one cell lineage, the radiate glial cells, serve as scaffolding and guiding lines for the movements of the emerging various cell types (see Rakic, 1981). Keeping their original attachments (or in some cases, e.g. cerebellar cortex only the outer attachment points) the original neighbourhood relations of the ancestors of most cells are essentially maintained through any subsequent twists and turns of the medullary tube or the brain vesicles and distortions caused by differential multiplication and growth of the inner (ventricular cell matrix) as well as by successive outward movement of later generated cells. It is hence quite logical that virtually all cells of the mature neural centres are derived—within a relatively narrow range in tangential direction (say 200–300 μm)—from the same ventricular epithelium of the early neural tube. Most central organs being several times thicker than 200–300 μm, the consequence of this will be that, especially in the cerebral cortex, all cells of common origin are organized in vertical columns. This basic developmental mechanism suffers various modifications in the details—that cannot be discussed here—and consequently the modular architectonics appear in different apparently radical variations (see Szentagothai, 1981, 1983), although the basic developmental mechanism is identical.

Figure 4 summarizes the main known facts of cerebral cortex organization. The upper part of the diagram demonstrates the so-called ‘columnar’ organization principle by showing two coronal (transverse) discs of the two hemispheres. The lower disc has to be understood cutting through the anterior part of the brain, while the upper disc has to be envisaged as being on a level behind the first. The section in front illustrates the connections between columnar modules in the same coronal plane, both with other modules of the same hemisphere and—over the corpus callosum, connecting the two hemispheres—with columns of the opposite. While such cortico-cortical connexions directed to ipsilateral cortical regions arise exclusively from the upper or outer (so
Self-organization in the nervous system

called supragranular) three cell layers (exclusively from pyramidal cells) of the cortex (in large majority from the third pyramidal cell layer), the commissural (callosal) fibres directed to the opposite hemispheres may arise from pyramidal cells of any layer II–VI of the cortex (also from 'fusiform' cells of layer VI). Pyramidal and other efferent cells of the (inner) infragranular cell layers (laminae V and VI) are projecting primarily to non-cortical targets (subcortical nuclei, brain stem, and spinal cord). Their callosal connections—if they have such—are not the main axons, but simply major side branches of the main descending fibers. The transverse section through the rear part of the brain (upper disc in the diagram) illustrates another important principle of cortical connectivity (Goldman-Rakic, 1984). Cortico-cortical projections of different origins to the same larger area of the cortex do often not converge upon the same colunar regions, but upon systematically alternating modules that show little if any direct overlap (i.e. no sharing of or invading each others spaces). Projections terminate more often than not, in longitudinal stripes (trenches of cortical tissue, if viewed in three dimensions) which are columns only in one diameter (having vertical margins) and series of columns in the perpendicular direction. As can be demonstrated by computer reconstruction from section series of labelled projections, such projections look more like the plastic maps of narrow mountain ranges, in which the individual fibres (or fibre groups) terminating in one column represent mountain peaks and the whole projection the mountain range.

The lower part of Figure 4 shows one individual columnar module stereodiarographically (adapted from earlier similar illustrations; Szentágothai 1987). The original module corresponds to a vertically oriented cylinder of 200–300μm diameter cutting through the entire depth of the cortex (3 mm in the human, less in most subhuman mammals). The module is primarily defined by one (in reality a group of several tens of) cortico-cortical afferent(s) placed into the centre of the cylinder. Such cortico-cortical afferents establish synapses in virtually all layers of the cortex in the tangential range of 200–300μm diameter. It is only lamina I on the surface where the terminal branches become more elongated and here its range becomes enlarged two to three times.

Fig. 4. Architectural principle of the cerebral cortex. Upper part of the diagram shows that the cortex can be envisaged as a mosaic of vertical columns, defined by the convergence of a group of cortico-cortical afferents. Only pyramidal cells are illustrated from which those situated in the outer three layers of the cortex project exclusively to other parts of the cortex, where the fibres terminate as cortico-cortical afferents. The two rows of columns illustrated in this diagram represent two coronal slices through the cortex to show the principle of systematic alternation in projections from different sources. Pyramid cells of the lower two layers project mainly to subcortical targets, but their collaterals, or some of the main branches, reach the cortex of the opposite hemisphere. Lower part of the diagram gives a radically simplified diagram of the cortical column: a vertical cylinder containing about 5000 nerve cells. ~ 60% pyramidal (Py. projective = output cells), ~ 40% are interneurons of various kinds, 20% of the total are now known to be inhibitory. The cylindrical column is defined anatomically by the cortico-cortical afferents (cortico-cort.) placed onto the axis of the cylinder. The specific sensory afferents (sens. aff.) are dominant only in the primary sensory areas, but secondary and further sensory areas have a similar input into lamina IV (lamination in Roman numerals at left margin) from the primary areas. The further local transmission of the specific input is indicated only by arrows (in outline if excitatory, and black if inhibitory). Two rows of large basket cells (Ba) at the border of layers III–IV and IV–V exercise inhibition over distances of up to 2–3 mm. This horizontal inhibition is assumed to narrow down the waist of the column in layers III–V (dark horizontal arrows), while excitation is conjectured to spread in radial direction both in lamina I and VI, either by the long terminal branches of cortico-cortical afferents in lamina I and VI. This distorts the original cylinder dynamically into the shape of an hourglass. Modified from Szentágothai (1987)
Other afferents—for example specific sensory afferents in sensory cortical regions—show a specific preference for lamina IV, but otherwise the modular (columnar) principle is the same. The difference between cortico-cortical and specific sensory (subcortical projections is only that while the cortico-cortical projections reach and establish synapses directly in all layers of the cortex, the specific afferent (sensory) input is directed specifically to lamina IV cells (with certain variations in the different sensory systems). However, due to a group of specific excitatory relay cells (so called spiny stellate cells) the afferent input pattern is immediately transformed into a vertically ascending (and descending), hence vertical columnar, pattern of excitation. Then the difference is only one additional intercalated neuron (and some additional recombination by convergence and divergence) in the sensory afferent projection.

The major complication in the cortical neuron networks is introduced by a couple of ten different types of inhibitory interneurons. A single representative type of inhibitory interneuron (a so called basket cell) is introduced into the diagram in full black (while excitatory cells and connexions are drawn in outlines). Many of the excitatory and most of the inhibitory interneurons (with one single very notable exception; for the details see reviews of Szentágothai, 1983, 1987) are not too specifically addressed to any cell type or any specific locus where they establish their synapses. Hence there is ample freedom (or quasi-randomness) in connectivity on the modular range (i.e. in tangential distance of 200–300 μm). While addressing is remarkably specific and precise on the macroscopic scale, i.e. in the longer cortico-cortical and subcortico-cortical connections. Columnar modules are remarkably similar built from mouse to man; the same cell types and kinds of synapses do occur on all phylogenetic levels. Even the cell numbers per column are quite similar (~5000 neurons). The difference is in the number of individual modules: a couple of hundreds or thousands in rodents and $2 \times 10^6$ in man. More columns mean, of course, many more connections; if each column has connections with, say, 50–100 other columns (and some additional subcortical sites), the difference between, say, rodent and man becomes apparent. More connections, of course, need more tissue volume (the cell size does not show any apparent correlation with phylogenetic level, body size, etc.), hence the cortical volume in man is about 3–4 times that of cortical volume in apes, while the number of cortical cells in apes is only slightly less than one half of that in man.

The original cylindric module is dynamically distorted into the shape of an hour-glass (rotation hyperboloid) by the preponderance of wide-spread excitatory connections in layers I and VI and by strong tangential inhibitory connections in the middle layers II–V. This is illustrated in the lower diagram of Figure 4, showing the widening of the original cylinder both at the outer and the inner surface of the cortex. Inhibition, conversely, tends to narrow the waist of the cylinder by corsetting the column around its middle. Many other preponderant connections operate within each module: there is much vertically arranged mutual inhibition and excitation between the upper and lower layers of the cortex.

Certainly, cortical anatomy shows many local mutual connections both inside the module and between neighbouring modules with very little of rigidly predetermined
paths—just what one would need for dynamic (self-organizing) interplay between neurons in the local range. Conversely, connectivity on the distant scale (for the total brain) appears to be very highly ordered and very specific for the species—again what one would have logically to assume.

The same principles hold essentially for all other parts of the central nervous system (see review of Szentágothai, 1983).

4. Pattern formation and pattern recognition

The term 'pattern' is one of the most popular expressions in present-day science. Biological pattern-generating mechanisms have quite similar structure, since they are frequently based on the notion of selection. The existence of selection mechanisms presupposes the operation of internal diversity acting on population.

Selection mechanisms occur at different hierarchical levels of the nervous system. On the lower levels, the regulation of the synaptic efficacy based on 'receptor desensitization' (Heidman & Changeux, 1982) can be described in terms of molecular interactions. Certain properties of single neurons, as selectivity for orientation (Blakemore & Cooper 1970), motion (Olson & Pettigrew, 1974) and disparity (Pettigrew), 1974) are modifiable by the patterns of stimulation they receive.

Our ideas on the formation of synaptic connectivity patterns are undergoing fundamental changes presently. According to Sperry's chemoaffinity hypothesis (Sperry, 1943, 1963) there would be a point-to-point projection between retina and tectum: order and orientation of topographic maps were derived exclusively from the chemical interactions between pre- and postsynaptic cells. In the original version of the hypothesis each retina ganglion cell was supposed to be labelled with a unique chemical marker, and the complementary marker on the tectal neuron was assumed to be recognized by some molecular recognition mechanism. Newer versions use considerably fewer markers.

The discovery of the so-called cell adhesion molecules (CAM) has now rendered the old discussion around specificity obsolete (see e.g. Edelman, 1984). Morphogenetic movements, including the outgrowth and the establishment of contacts by nerve processes have to be envisaged as dynamic processes, starting with a small repertoire of CAMs that becomes more and more differentiated by some epigenetic mechanism until it reaches a relatively stable 'final' stage on cell-to-cell contacts. The selective stabilization hypothesis posed already before the discovery of CAMs, (Changeux, Courrage & Danchin, 1973; Changeux & Danchin, 1976; Changeux Heidmann & Patte, 1984), offers a gene-saving mechanism for specifying ordered neural structures. While Changeux considers the spontaneous activity of neurons, as the units of selection, the other approach (Edelman, 1978; Edelman & Finkel, 1984; Edelman, 1987) emphasizes the role of groups. The stages of selection of neural patterns are the group confinement (i.e. the limitation of group size), group selection and group competition. Both theories try to unify the microscopic-molecular aspects of synaptogenesis and the macroscopic mechanism of formation of global mappings.

Studying the mechanisms of the formation and plastic behaviour of ordered neural structures the neurobiological relevance of the concept of 'pattern formation by learning' can be illustrated. Model studies suggest that topographic mappings are not genetically determined, but might be generated by (although perhaps genetically programmed) self-organizing mechanisms (Amari, 1983, Kohonen, 1984; Erdi & Barna, 1984; Bienenstock, 1985; Cottrell & Fort, 1986; Ritten & Schulten, 1986). According to our simulation results, randomness has a positive role in the formation of ordered neural structures:
both the retinotectal connections and also ocular dominance (Barna & Erdi, 1986; Erdi & Barna, 1988) columns might be generated by noise-induced (and activity-dependent) self-organizing mechanisms.

The pattern formation of synaptic connections can be associated with ontogenetic development, plasticity, memory and learning. Models of all these phenomena are based on the original or modified Hebb's (1949) association rule. However, Anti-Hebb's rule (e.g. Palm, 1982; Barlow & Foldiak, 1989) could be adopted for describing 'decorrelation' phenomena.

The understanding of the mechanism of the invariant pattern recognition is a great challenge for theoretical neurobiologists. The brain has to organize the representation of objects, providing for invariants concerning size, position and orientation. The classical algorithm of pattern recognition based on modifiable (synaptic) connections for classifying certain sets of patterns as similar or distinct is the reputed 'perceptron' (Rosenblatt, 1962). Modifying the perceptron concept Fukushima in a series of papers (e.g. Fukushima, 1984) suggested a multilayer neural network model, which has the ability of rapid self-organization. The main problem with Fukushima's approach is that the invariance is explicitly wired. Relation concepts of pattern recognition (e.g. Glunder, 1986a, 1986b; von der Malsburg & Bienenstock, 1987; Bienenstock & von der Malsburg, 1987) are based on inner pattern relations.

Biological information processing, and in particular pattern recognition can be treated as a hermeneutic process, emphasizing the importance of 'preunderstanding' (Vorverstàndnis). Accordingly, pattern recognition is a cyclic negative feedback process reconciling the interpretation of the whole and the preunderstanding of the parts.

Kohonen (1984, 1986) proposed an algorithm for the self-organizing mechanism of feature-detecting functions related to external input: ' . . . it has been almost a mystery how a physical system can automatically extract various kinds of abstraction from the huge number of vague sensory signals . . . '.

We cannot avoid mentioning the new concepts of pattern recognition evoked by recent developments of statistical physics. Global dynamic behaviour of nonlinear networks of elements, which can be considered both as neurons and as computing structures, can be treated by statistical methods (e.g. Choi & Huberman, 1983; Hopfield, 1982; Hogg & Huberman, 1985, Hopfield & Tank, 1985). The Hopfield model of associative memory is appropriate for describing 'learning' and 'pattern recognition' in spin glasses (Kinzel, 1985).

The overwhelming majority of mathematical techniques for invariant recognition applied in technical systems has no neural implementation, for what they are often criticized. Speaking about 'biological relevance' we agree with Hochberg (1987): 'Machines should not see as people do, but must know how people see'.

5. The essence of neural

What we tried to illustrate in the previous sections is that the concept of self-organization is useful for a better understanding of neural systems, and it might be a fundamental building block of any future (would-be) general brain theory.

Neuroscience itself is changing from being a descriptive science to becoming an axiomatic (although not physics style) discipline. While 'bottom-up' approach starts from neural elements of known functions and studies the behaviour of more complex systems, 'top-down' analysis tries to determine the neural implementation of known functional behaviour. The change of an operative brain theory depends on the interaction of these two approaches.
Self-organization in the nervous system

Selectionist brain theories (Changeux, 1985; Edelman, 1987) and Grossberg’s (1988) adaptive resonance theory are not based on but at least partially apply (explicitly or implicitly) the concept of self-organization.

Changeux’s approach emerged from biochemical studies on cholinergic muscle innervation. The selection stabilization hypothesis, however, has been generalized (perhaps a little bit overgeneralized) for offering a mechanism of storing and perceiving mental objects. The basic concept of the theory is that the brain spontaneously generates transient mental objects, prerepresentations that exist before the interaction with the environment. Some resonance may occur between the internal presentation and the ‘image’ of the external object (Heidman, Heidman & Changeux, 1984).

From all the brain theories presently available the consequently selectionist hypothesis of Edelman (1978, 1987; Edelman & Finkel, 1984) is an attractive concept, already by virtue of its intellectual challenge. However, it is exactly here, where its author meets his hubris (in the sense of the ancient Greeks) when offering predictions about what observations would falsify the theory. For example: it would be proven false, if individual nerve cells were found in the cerebral cortex that would specifically signalize concrete complex visual patterns, that are of ecological significance for the animal. This is exactly what is being claimed by some experimental physiologists. We would attribute less importance to such contradictions which could be explained away one way or another, but are more concerned about some of the unlikely consequences, into which any radically and exclusively selectionist theory has to lead necessarily. Apart from difficulties from the exceedingly brief somatic lifetimes in numerous invertebrate forms, leaving little if any room for explaining very complex forms of behaviour, our misgivings arise on the other end of the behavioural scale, with Edelman’s flat rejection of the reasoning of Noam Chomsky (1980; Lumsden & Wilson, 1981). As we see it, the deep logical structures of the specifically human ability of speech acquisition are wired genetically into the upper surface of the human left (to much lesser degree into the right) temporal lobe, exactly as the neuronal mechanisms for moving a limb are wired (Szekely & Czeh, 1971) into the limb-innervating segments of the medullary tube.

The adaptive resonance theory of Grossberg (1988) has been applied also to speech and language coding (see particularly Cohen & Grossberg, 1986). Complex speech and language units emerge as the result of self-organizing multicellular interactions.

The essence of the neural organization is—at its very basis and its ultimate origin—self-organization of spontaneous (random) neural noise into spatiotemporal patterns of activity. Everything else is secondary: viz. establishment of connections with the environment over receptors, epigenetic selection and selective stabilization of appropriate functional chains of genetically not preprogrammed neuron connections, and gradual integration with the primary activity patterns of various kinds of flows of information [including eventually those carrying mental (psychic, cognitive, etc.) phenomena]. The idea that we are trying to bring across is but to show that the old paradigm of the neural as a ‘reflex machine’ is hopelessly outdated and has to yield to the new paradigm being now built around the central concept of self-organization.

References


