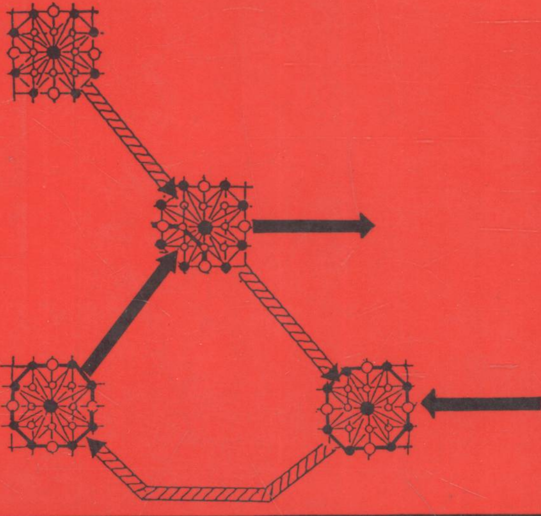
AN ADAPTIVE NEURAL NETWORK
the cerebral cortex

Y. BURNOD



TP18
B966
9260333

AN ADAPTIVE NEURAL NETWORK

the cerebral cortex

YVES BURNOD
INSERM Researcher
at the Institut Pasteur

E9260333

图书馆藏书

MASSON PARIS | 1990 | PRENTICE HALL LONDON

The Masson edition of this book is published in the series *Biologie théorique* edited by G. CHAUVET and H. LE GUYADER.

Published in 1988 in France, Italy, Portugal, Spain, Switzerland and Belgium by Masson, and copublished 1990 by Prentice Hall International for distribution throughout the rest of the world.

© Masson, Paris, 1988, 1990

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission, in writing, from the publisher.

ISBN 0-13-019464-6

Printed in the Republic of Singapore

Library of Congress Cataloging-in-Publication Data
British Library Cataloguing in Publication Data
are available from the publisher.

AN ADAPTIVE NEURAL NETWORK

the cerebral cortex

Contents

图书馆藏书

此为试读，需要完整PDF请访问：www.ertongbook.com

Introduction

Tool-making and language: these capacities that have so dramatically multiplied the adaptiveness of the human species, issue in some poorly-understood way from the human brain.

It is generally recognized that the cerebral cortex, extensively developed in man, plays a crucially important role in the elaboration of these capacities. We are gradually acquiring an understanding of some of the fundamental aspects of cortical structure and function (Mountcastle 1978), and we can now attempt to relate these mechanisms to higher-level human capacities.

In the investigation of these complex phenomena, two general research strategies have greatly contributed to our understanding of cerebral mechanisms:

1. Experimental results from neurobiology and psychology. Powerful experimental methods, proceeding by testing verifiable hypotheses, confronting them with precise observations of concrete reality, have furnished a wealth of information. But this information has been gathered in heterogeneous experimental contexts and these facts need to be fitted into an overall theoretical structure in order to adequately explain the complexities of human adaptation.

2. Artificial intelligence has led to the construction of systems whose information processing capacities approach those of human beings. In particular, adaptive automata networks which perform parallel distributed processing have many analogies with mental processes like learning and generalizing from examples. But these artificial systems are still far from the adaptive wealth of the human brain.

The aim of this book is to propose a theoretical model of the cerebral cortex with an approach combining Neurobiology and Artificial Intelligence. This model defines a specific cortical function, based on the cellular organization of the cortex and shows how this cortical function can explain the principal adaptations and forms of learning possible to the human species: recognition of complex forms, visually-guided hand movements, execution of structured programs, language learning. The theory does not consider all possible aspects of the cortex but focuses on those adaptive functions which are already partially solved by Artificial Intelligence.

Real neural networks have inspired several theories of adaptive automata networks (see review in Rumelhart 1986). In these theories, every behavioural or cognitive process resulted from the coordination of a large number of cells or cell assemblies (Hebb 1949), each roughly localized in different regions of the brain, but all working together in dynamic interactions (Luria 1965).

These theories assume that information processing is parallel and distributed over wide sets of interacting elements - or automata- each sending excitatory or inhibitory signals to other units (Rumelhart 1986)).

The properties of the processing elements were inspired by the basic properties of the neural hardware. Biological neurons were modeled as logical decision elements described by all-or-none state variable (Mc Cullogh and Pitts, 1943). Adaptive properties of these networks were based on local "learning rules" for each automaton that can behave as single unit analogs of associative conditioning (Hebb 1949): this local learning mechanism adjusts the strength of connections between units based on information locally available at the connection, for example when input and output are simultaneously active ("Hebb rule"). Learning rules of processing units can be mathematically determined in order to produce a global adaptive function for the network (Widrow, 1960). In these models, behavioral learning is a result of tuning the connections: the memory and the knowledge is in the strength of connections.

These neural-like networks have many interesting properties. They can learn to associate representations, for example the association between the visual form of a word and its meaning . Although far from real neural networks, these parallel distributed models can account in detail for psychological data on the processes of pattern recognition, speech perception and generalization of learning from examples. As with human memory, the network tends to retrieve what is common to different input patterns. Such distributed mechanisms can learn rules from examples. For instance, a network can learn rules to construct the past tense of verbs from their root forms; in this case, the behaviour of the system follows the same learning stages as children, such as transient regularization of irregular verbs.

These automata networks can also precisely describe processes at the neurobiological level, like self organization of cortical maps (Kohonen 1984) or specific information processing performed by a cortical area; it is then possible to compare the behaviour of automata with the properties of cortical neurons experimentally analyzed in these areas (Zipser 1988).

A) We propose a model architecture for the four main organizational levels of the cortical system.

Our basic conception shares common features with these models which are neurally inspired: information processing is parallel and distributed, and learning is produced by changes in the connection strengths between elements.

But our approach is quite different: its originality is to start from the physical reality of the nervous system, with present available knowledge, to propose a network of automata specific to the cerebral cortex and capable of generating similar adaptive functions such as invariant recognition, manipulation and language.

Starting from the physical reality means taking into account the different levels of the cortical system: cells, columns, maps and areas; each level gives specific adaptive properties that are not presently included in models.

We focus on the constraints which arise from the physical reality at the different levels of the brain. We describe a biological system which has an adaptive efficiency that comes from a variety of means and levels more than a mathematical efficiency that comes from a single formula for a local rule resulting in a global function performed by the network.

This is because the adaptive properties of a biological system depend not on its organization at any one level, but rather at different hierarchical levels, each lower level being tightly articulated into the next highest level: an isolated cell can execute basic adaptations such as recognition and orientation towards a goal; cellular tissues and organs, by their three-dimensional geometry, facilitate certain exchanges with the outside world; at a higher level, the different organs co-operate to ensure a global energy balance for the whole organism.

In order to fully explain the adaptive characteristics of a biological system like the cortex, we are obliged to consider all these levels-- none can be omitted. It should be clear that one cannot arrive at explanations for behaviour by considering only the cellular level; even if it is possible to show that memorization is a general property of the nervous cells, behavioural learning capacities are not: learning in humans and in invertebrates is quite different. On the other hand, it is difficult to fully explain the adaptive capacities of the human brain by considering only the interactions between the different cortical areas, without knowing what the cortical cells are doing.

Our theory is original in that it proposes a model architecture and functioning rules as compatible as possible with experimental knowledge at the four main organizational levels of the cerebral cortex:

1. Cortical cells - pyramidal neurons and interneurons- share many common features with neurons in other structures but they have specific properties of integration and memorization of input information organized in layers.

2. Cortical "columns" are groups of cooperating cells. If we imagine the cortex as a two-dimensional tissue, with input and output at the bottom, a column forms a circuit running up and then down perpendicular to the surface of the tissue; the entire cortex itself is made up of billions of such columns. These columnar circuits perform transformation of the inputs, receiving and passing in the process signals from or to other such columns. This input-output processing - the "columnar function" - have adaptive and learning properties which do not exist at the single cell level.

3. Cortical "maps" are formed by contiguous sets of columns which have ordered topological input-output relations with sensory or motor areas, or with other nervous structures. At this "regional" level, the cortical network integrate columnar processes into behavioral adaptations such as positioning or recognition.

4. The network between cortical areas has an overall architecture which organizes the main information flows, such as relations between audition and phonation or between vision and manipulation. But this network can also integrate successive learning experiences in a coherent functional system such as language with its different components, auditory and phonetic, syntactic and semantic.

A number of observations will serve to orient the approach that we will undertake.

a) Cellular level: basic integrative properties of neurons.

Although the cortex has often been described as a general integration and learning system, these two capacities are not exclusively cortical in nature, but seem to be general neuronal properties since they exist in species such as invertebrates that have no cortex (Kandel, 1977b). The property of "memorization" seems to be generally found in a wide variety of nervous systems: this memorization has been studied at the cellular level in invertebrates (Kandel 1977b; Alkon 1982) and, in vertebrate nervous systems, in neural structures with a regular geometry, such as the hippocampus (Bliss 1973; Andersen 1977) or the cerebellum (Ito 1982).

b) Modular level: columnar processing.

The cortex has been subdivided into a number of specific areas, each of which has a functional specialization (Luria 1973), such as: motor sequence, spatial positioning, visual analysis, production and comprehension of language, abstraction, etc. But in spite of this functional diversity, detailed analysis at the cellular level has led researchers to formulate a structural and functional principle generalizable to the entire cortex: the "cortical column" (Mountcastle 1959; Szentagothai 1975).

The functional characteristics of this "column" have been detailed in the cortical receptive (Hubel 1977), motor (Evarts 1974), associative (Mountcastle 1975), and frontal (Fuster 1973) areas. It is therefore important to understand how such columns, functioning according to general principles of cellular interactions, accomplish the informational transformations typical of human learning.

This basic function is not a simple associative mechanism: for example language production demands the activation of a highly specialized mechanism that cannot be reduced to simple sensory-motor conditioning (Chomsky 1970). In order to define this specific operation, it is important to note that the internal structure of human cortical tissue is in direct evolutionary continuity with the cortex of mammals and the other primates (Cajal 1911). Consequently, it should be possible to find, in these species, a cortical capacity that is the functional precursor of the human cognitive capacities.

This basic cortical operation should be explained from its cellular components. In theory it is possible to deduce some properties of a neural tissue from neuronal shapes (Rall 1964), ionic channels (Traub) and long term changes of synaptic transmission (Hebb 1949). Such reconstructions have been attempted, for example for the cerebellum (Marr 1969, Pellionisz 1982) and the cerebral cortex (Marr 1970, Eccles 1981), but these models have not considered a specific columnar operation.

c) Regional or "tissular" level: construction of basic behavioural functions.

We now need to understand how the basic columnar function common to all the cortical areas, can generate extremely varied behavioural adaptations, depending upon the links of the cortical columns with sensory or motor organs and upon the cortical network which relates cortical maps. In order that we may understand the progressive learning of specific global behavioural functions, such as recognition, or positioning in space, we must find a model for this cortical network; its geometrical features may induce many combinatorial properties which determine learning capacities. For example visual guidance of arm movements is easier with a network which forms all possible combinations between command of muscles and regions of the visual field; verbal imitation of words is highly favoured by a network that links all positions of the vocal apparatus with auditory frequencies.

d) Global level: integrated learning.

The cognitive development in the child is an ordered process (Piaget 1968). Cortical maturation in the human species is slow, particularly in the frontal lobes which are much more extended than in other species. This neurobiological basis may be an important factor of the cognitive development. The construction of mental

此为试读，需要完整PDF请访问：www.ertongbook.com

images and words should be highly dependent upon the growth and the resulting geometry of the cortical network: it is important to see how the semantic structure of words and the syntactic organization of language can be learned with this network.

B) We propose a "cortical mechanism" that can explain the adaptive capacities of the cortex.

In our model we focus on the behaviour of the columnar automaton that is the behaviour of a group of cells composed of pyramidal neurons and different types of interneurons tightly linked in a cylinder perpendicular to the cortical surface: we consider the columnar inputs as the combination of all inputs to each cell of the column which originate externally from the column, the columnar state to be the combination of the states of each cell and the columnar output, the combination of external outputs from each cell. Furthermore, just as it is possible to describe the process of memorization for a cell, by which it changes its behaviour, so this idea can be applied for columns, where, in learning situations, the columnar function may be altered.

Compared with a cellular automaton, a columnar automaton has more functional and learning capacities. To take an analogy with computers, we could imagine a cell to correspond to one computer memory location, and a column as an area of computer memory in which may be stored different computer programs to perform different subfunctions. The memorization of a subfunction by a cortical column can be compared with the development of a small computer program.

Furthermore, just as a large number of computer programs are integrated to provide a complex system, so cortical columns, each with its particular subfunction, are assembled by learning to provide a far more complex behavioral function. Learning of behaviour by a cortical region could be compared with the step by step integration of small programs to form a computer system. The major difference with computer programs is that columnar subfunctions possess their own intrinsic rules for building up larger structures.

In chapter II, we define this columnar automaton by an in-out table which determines the two outputs of a column (intra and extracortical) from its inputs (cortical and thalamic) and from its previous state. It is based on neuronal processing performed by the different cell types of the column and integrates known physiological properties such as gating by cortical inputs, lateral inhibition and vertical amplification. We define two learning rules that enable sets of columns to be assembled in integrated mechanism, one in the top-down direction from goals, and the other one in the bottom-up direction from the external information. We base these learning properties upon the memorizing properties of different cortical cells (pyramidal neurons and five interneuronal types) with a logic based upon cell position in the network.

This model of "columnar automaton" meets three essential requirements:

1. The ubiquity of the columnar architecture implies that the operational mechanism is the same in any region of the cortex. This columnar architecture provides a common framework upon which we can generalize; variations of cellular texture in different areas throughout the cortex ("cytoarchitectonics") are included in the variations of parameters of the columnar operations : they correspond to local adjustments of a common function.

2. The proposed mechanism is specific to the cerebral cortex in that it can only be deduced from the unique cellular structure of the column, but not from other brain textures such as the hippocampus or the cerebellum; this cortical process is basically an active "searching mechanism" that provides the possible pathways in order to reach a goal from any initial position. This mechanism which is compatible with experiments that relate cell activity and behaviour, can explain two main cortical properties, goal-directed behaviour and active self-driven learning.

3. This unique process can construct diverse behavioural functions depending upon the connectivity of each cortical area. It provides a unitary explanation for apparently diverse cortical functions such as visual guidance of hand movements, learning by imitation or the symbolic manipulation of language. In chapter IV, we examine the construction of these diverse functions depending upon the network of connections of each cortical area and upon its maturation. For this purpose, we propose in Chapter III an integrated model of the network between cortical regions, based on the present anatomical knowledge, in order to focus on its geometrical and combinatorial properties.

Each specialized region is formed by billions of columns which all represent possible actions, goals and subgoals; columnar automata construct the possible pathways in order to reach such goals from any initial position determined by environmental situations;

- In parietal areas, these call trees perform visual guidance of hand movements to reach a target, whatever the initial positions.
- In temporal areas they are algorithms that can recognize forms independently of their size or retinal position.
- In frontal regions which have short term memory properties call trees organize "structured sequences" with a variable number of internal levels such as sequences of words in language.

The network between columns and between cortical areas has a precise architecture that influences the learning capacities of the cerebral cortex. In the last chapter we will see how this architecture plays a key role in the cognitive development of the child: the external world will be represented in the child's brain in a coherent way.

C) The mechanism that we propose is compatible with the experimental results at all four levels of cortical organization.

Models of cortical function hitherto proposed are not equivalent when one accumulates constraints. We adopt a strategy that combines the neurobiological and artificial intelligence approaches and our model attempts to satisfy two criteria:

- functionality: the proposed mechanism produces behavioural functions and has adaptive capacities;
- compatibility: the model is consistent with what is known of the cellular structure of the cortex and the behavioural capacities attributed to it.

Many experimental approaches and techniques have been used to study the various levels of cortical organization, such as biochemistry, histology, anatomy, electrophysiology, neuropsychology, experimental psychology, and cognitive psychology.

The results are so extensive that it is impossible to take them all into account: our strategy has been to try to consider results from a variety of sources.

One of the cardinal requirements of the kind of experimental compatibility that we seek is that the model takes into account the different dimensions and the different levels of the cortical system. At each level we try to stick to general principles experimentally established in order to have a good representation of the immense amount of existing material that describes the path from cell function to language production. This work is not an attempt at a complete explanation for language, it is a description of a neuronal system that can generate language with its different syntactic and semantic components.

For our purposes it is necessary to show that the theory makes sense in terms of basic concepts before going into finer details. For example, autoproduction of language by children is a fundamental phenomenon which is not explained by simple associative models (Chomsky); furthermore, there are successive stages that are the same for all children and that do not depend on their specific environment as would be predicted by simple associationist schemes (Piaget). We do not want to examine supporting data in detail (other people have done this) and we do not wish to enter into discussions such as the precise description of stages of child development. We merely try to show that autoproduction of language in an ordered way is a direct consequence of cortical properties. The main difficulty is to discuss behaviour and language processes with concepts derived from biology. All the different concepts addressed in the last chapter on cognitive processes (cortical image, word, etc.) are related to a precise, imitable, neural process.

The different parameters of cortical organization are far from being completely known. The partial knowledge that we have of synaptic function and the action of neuro-mediators at the cortical level is the fruit of many experiments; we have tried to integrate the principle characteristics.

Different models of neuronal function have been elaborated from experiments performed on systems other than the cerebral cortex. We have chosen for study those that appeared to be best supported by experimental evidence, and from these models, considered to be general, we have derived rules that were subsequently applied to the parameters of the cortical network.

It is clear that the kind of mechanism attributed to the cortex here goes considerably beyond what is factually known, and enters the realm of the hypothetical. But this model makes predictions which are experimentally verifiable; for instance, electrophysiological techniques that permit the quantification of neuronal activity could confirm whether or not the cells follow the logic that we propose. Interactions of large populations of neurons cannot be directly analyzed as yet, but behavioural consequences of these interactions can be tested.

D) We adopt a systematic approach in order to focus on the coherence of the cortical system which is constructed from a single cell and produces a global adaptive behaviour.

Above all, the cerebral cortex is a part of a multi-level biological system which has a global coherence.

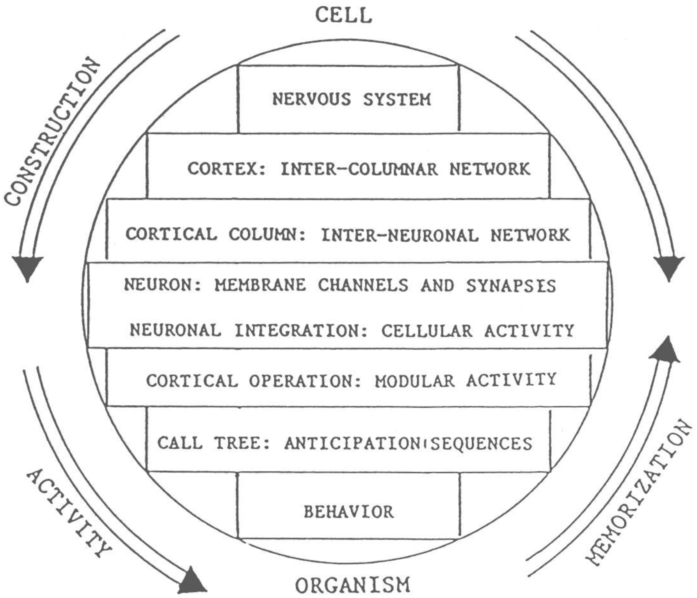


Figure 0-1 . From cell to behaviour: global coherence of the cortical system.

FIGURE 0-1 depicts the basic architecture of a multicellular organism and illustrates its fundamental property: the whole system is constructed from a single cell (upper part, downward arrows), and the results of its cellular interactions produce a globally adaptive behaviour (lower part, downward arrows).

Adaptive behaviour is possible thanks to two conditions:

- the cellular groups, such as organs, represent and amplify the different basic functions of the cell: energy, development, reproduction, communication.

- neural and hormonal communications coordinate the complementary functions of organs to ensure the energy balance of each individual cell, and therefore the survival of the organism in its environment.

The nervous system can be diversified within the limits of these two constraining conditions (fundamental programs and energy balance). Different parts of the nervous system (retina, spinal cord, cerebral cortex, etc.) are thus specialized to control various interactions with the environment, which contribute to the survival and the reproduction of the organism: for example recognition, oriented movement or communication.

Such adaptive processes can be described from their end result. This desired end result, we will call the "goal". For example, in homeostatic mechanisms, goals are constant physiological parameters regardless of external variation. In another example, placing the hand on an object is a spatial adaptation, the goal is the minimal distance between the object and the hand. More generally, a behavioural adaptation corresponds to a goal which is a minimal or a maximal value of a parameter which represents an interaction with the environment. In visually guided hand movements, the parameter of the interaction would be the distance between the hand and the object. The goal is reached whatever the initial position by minimizing or "optimizing" this parameter by any combination of muscular contractions ; many different movements may have the same final result.

Whenever we refer in this text to the term "optimization", we mean the process of achieving an end result, which is defined as a maximal or a minimal value of a parameter, for example a minimal distance. Depending upon the parameter which is optimized, adaptations can be placing (minimal distances), recognitions (optimal matching to an external form), motor programs (optimal temporal patterns); communications and language correspond to more sophisticated forms of adaptations (optimization by social interaction) .

The cerebral cortex can generate all these specific forms of behavioural adaptations. Adaptive properties of a system like the cortex originate from three complementary factors: its structure, its function, and its learning capacities (shown by the three arrows of Figure 0-1).

a) STRUCTURE: Construction of a multilevel network of automata from a single cell.

The "structure" of a biological system is the organizational plan, the network of interactions between the cellular elements. Particularly important for a biological system is the progressive elaboration of complex structures from a single cell (ontogeny).

During embryogenesis, cellular divisions generate larger and larger groups of cells. Functional systems are genetically determined from the larger to the finer levels: organs, tissue, modules and cell types (downward arrow in Figure 0-1). The structure of the cortical network will not be described in an all-or-none manner, but as a result of a construction algorithm which determines connections between areas, between maps, between columns and between neuronal types.

This construction is based upon the universal cycle of cellular division accompanied by diverse cellular actions controlled by the genes: cell migration forms layers of neurons, membrane extension forms axons and dendrites, intercellular communication forms synapses, etc. Simple genomic mutations can thus produce a great variety of neural networks (phylogeny) and the universal division-mutation process creates new adaptive networks by a simple "trial and error" method.

b) FUNCTION: From neuronal processing to adaptive behaviour.

Associated with each element or automaton is a property which is its internal state and this internal state we refer to as the "activity" of the automaton. An element in a particular state of activity, when receiving inputs from other elements may change its current state of activity as well as the output it produces. This input-output operation can be described in mathematical terms: for example it can produce an oscillation (rhythmic pattern generator) that is important for movement, or "filter" a patterned stimulus, that is important for recognition. More generally, we have described behavioral adaptations as "optimizations" of parameters which represent interactions with the environment; such optimizations can be obtained by two complementary neural operations:

- the comparison between the current state of the system and the desired goal;

- a modification of this "distance" by feedback actions until the optimal value is attained.

Throughout the text when we refer to an action, we mean the idea of the result of the neural activity on a parameter which measures a behavioural adaptation, such as a movement that changes a distance, a feature extraction that may increase matching of an image with a stored pattern, or even a word that modifies a social relation.

c) PLASTICITY, LEARNING AND MEMORY: from long term cellular changes to learning capacities.

Neural networks can learn and memorize new adaptations. Description of learning and memory is viewed as long term changes of the interactions between the automata: between two elements of the network, the presence of a physical connection in itself does not necessarily lead to communication. Associated with each physical connection, is the concept of a "coupling coefficient" which modifies any signal passing along the physical connection, either to strengthen the signal or possibly to inhibit it altogether.

This coupling coefficient can itself be modified on a long term basis so that an element reacts differently to particular inputs, and produces new types of outputs. This process is referred to as "memorization" when it depends upon activity in the network. "Learning rules" describe how "coupling coefficients" are modified in relation with activity; long term changes depend upon critical "activity patterns" which are particular combinations of inputs and internal states.

Two aspects of the network can be identified. The first, the physical substratum, comprises sets of neurons along with their interconnections; the second aspect, the "functional network", is the result of the superimposition, upon the physical substratum, of the results of memorization stored in coupling coefficients by elements of the network, by which information is actually processed.

Patterns of connectivity between neurons are modified by experience (learning rules), and they can internally represent the properties of the external world. Adaptive mechanisms are amplified by memorization processes. New adaptations can be learned by continual modification of the coupling coefficients with activity patterns which depend upon the result with respect to the goal.

Different types of neuronal tissues will have different learning capacities, from very simple forms (pavlovian conditioning) to very elaborate forms (language learning). We shall try to demonstrate that these more elaborate forms are due to the specific properties of the cortical tissue. The rules that we propose for columnar automata do not make the cortex into a simple associative system, but instead the material support for a general mechanism of adaptation that includes both an active searching mechanism and a highly structured learning process. We will show how this fundamental mechanism can assemble new functional networks which produce not only adaptations like form recognition and placing responses, but also basic linguistic functions, which are specific expressions of a more general adaptation of communication.

We shall use a systematic approach of the adaptive capacities of the cerebral cortex and we shall proceed on three distinct parallel pathways:

1. We shall describe with simple rules structure, function and memorization properties at the four levels of the cortical system : cells, columns, maps and organs.

此为试读，需要完整PDF请访问：www.ertongbook.com