Chapter 13
Holographic Memory: A Novel Model of Information Processing by Neuronal Microcircuits

Alexey Redozubov

Abstract In the proposed model, each cortical minicolumn possesses a complete copy of the memory characteristic of the entire cortical zone to which it belongs. Hence, the cortex has holographic properties, where each fragment of an information carrier contains not just a part of the information but a complete copy. It is argued that each minicolumn encodes the new information using its own interpretation. Such transcoding is equivalent to considering the source information in a particular context. The model suggests that the cortex zone is a space of possible contexts for interpretation. The presence of a full copy of the memory at each minicolumn allows to determine which context is most suitable for interpreting the current information. Possible biological mechanisms are discussed that could implement the model components, including information processing algorithms that enable high computing power.

Keywords Holographic memory • Microcircuits • Information waves • Hippocampus • Meaning of information • Membrane receptors • Cluster of receptors • Cerebral cortex • Dendrites • Combination of neurotransmitters

13.1 The Propagation of Information Waves

13.1.1 Waves at Cellular Automaton

A cellular automaton (Von Neumann and Burks 1966) is a discrete model, which describes the regular lattice of cells, the possible states of the cells and the rules of changes between those states. Each cell can be in a finite number of states, for example, 0 or 1. For each cell, we define an area that contains its neighbors. The current state of the cell and the states of its neighbors determine the next state of the

A. Redozubov (✉)
St. Petersburg, Russia
e-mail: galdrd@gmail.com

I. Opris, M.F. Casanova (eds.), *The Physics of the Mind and Brain Disorders*,
Springer Series in Cognitive and Neural Systems 11,
https://doi.org/10.1007/978-3-319-29674-6_13

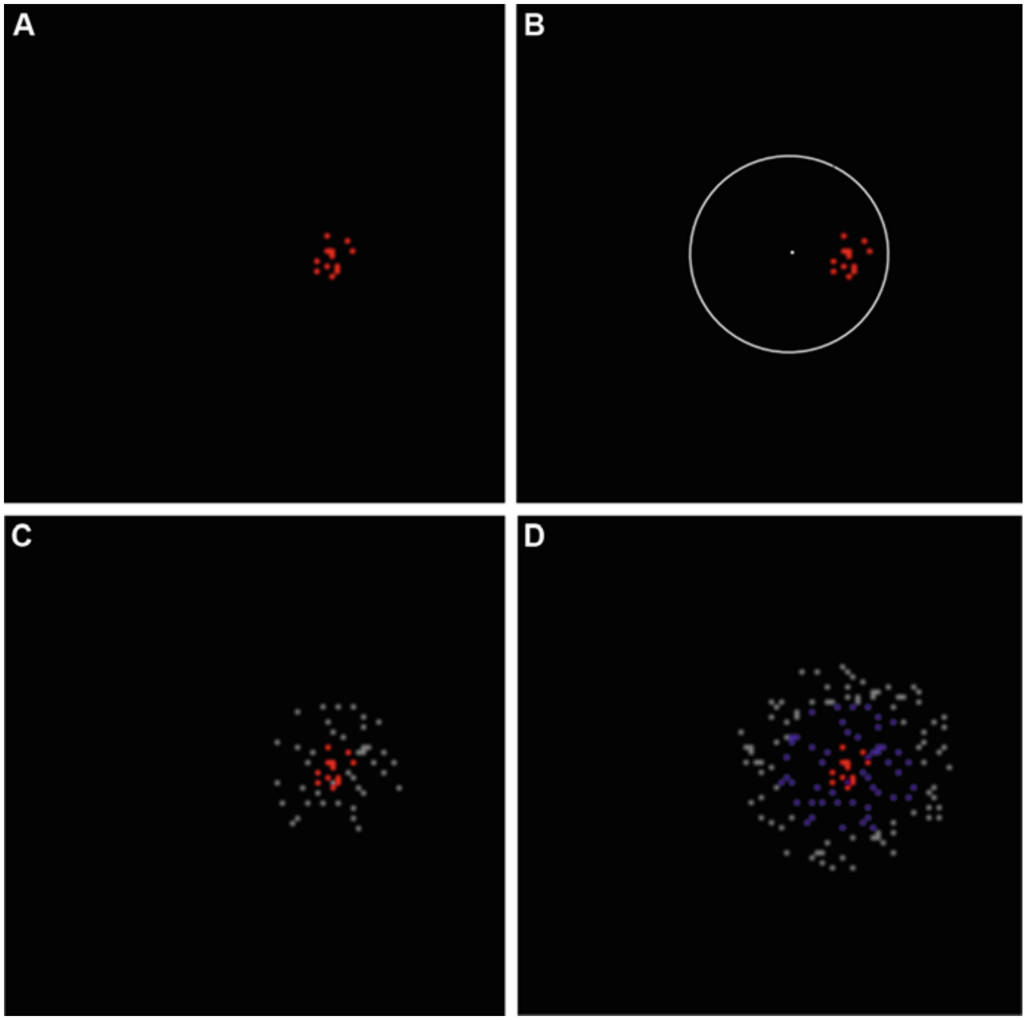

Fig. 13.1 Patterns of propagation. (**a**) the pattern of the initial activity. Only the active elements are shown. Elements are depicted tightly without a gap. Each pixel of the image corresponds to a single element. (**b**) the tracking field of an element and active elements within. (**c**) the first step of the simulation. The wave activity (*gray*) in front of the initial activity (*red*). (**d**) the second step of the simulation. The wave front propagation. Elements in the state of relaxation are painted in *blue*

cell. The most famous example of cellular automata is the "Life" game (Gardner 1970). Potentially, during the selection of a next state, cells can consider not only neighboring states and the transition rules, but their previous state changes too. In this case, we consider cellular automaton with memory.

Let us consider cellular automata with memory. Let's place its elements (automata cells) on a regular grid. For each element, we define its neighborhood that is the tracking zone of this element. Suppose that a compact pattern of activity appeared somehow on the automaton plane (Fig. 13.1a). The compactness of a pattern means here that all active elements fall in space within the size of tracking zone.

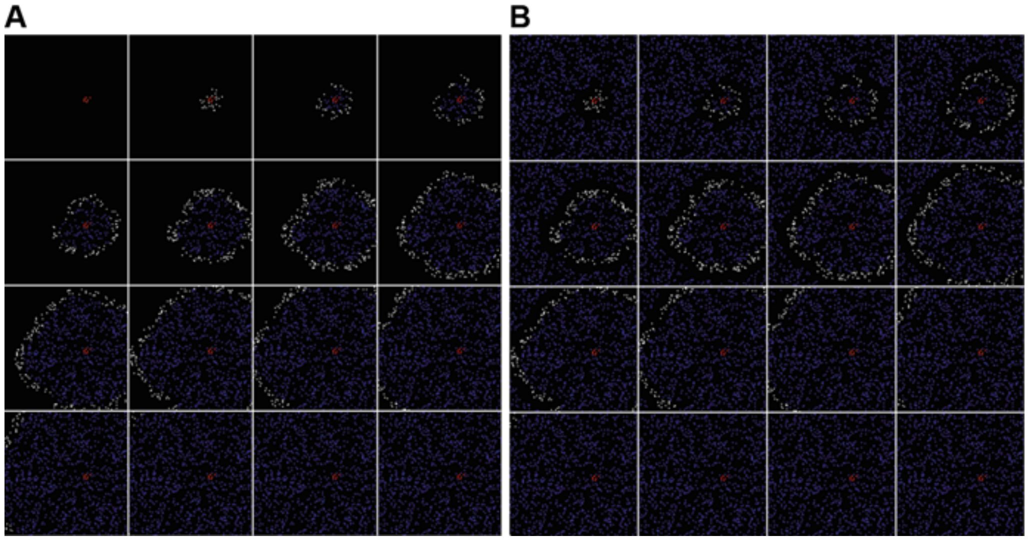

Fig. 13.2 (**a**) a series of initial cycles that propagate the wave activity pattern. *White dots* are the active elements forming the wave front. *Blue dots* are the elements in the state of relaxation against the spreading signal. (**b**) wave propagation on already trained automaton

Now we count how many active elements fall in the tracking field of each elements (Fig. 13.1b). Let's define some small probability p_{in} (roughly 3 % for the provided model). For each element in the quiet state, to count the number of active elements within the tracking field that exceeded a certain threshold, we perform the following procedure. Let's force an element to switch into an active state randomly, with the probability p_{in}. Accordingly, an element remains in an inactive state with probability $1 - p_{in}$. For that element, let's remember its choice and active elements in its tracking field. As the result of this procedure, a randomly generated pattern of activity forms around the pattern of initial activity (Fig. 13.1c). For reasons that will soon become clear let's define this emerged activity as wave-like. On the second step of the simulation, the elements located on the perimeter of the wave activity zone, will "observe" significant activity in their tracking fields. For those ones with the activity exceeded the threshold we repeat the previously described activation procedure. The elements activated on the previous step, using the parameter p, we transit to the state of relaxation. We deactivate them and block for a certain amount of time $\mathbf{T_{relax}}$ their ability to be activated by the pattern that caused their activity before (Fig. 13.1d).

By repeating the simulation steps, we get activity propagating across the automaton with a certain unique randomly generated pattern (Fig. 13.2a).

By changing the initial pattern, we scatter the wave front with its randomly generated internal pattern. Wherein the machine elements to be remembered what patterns have already ran through them. Due to this memory can be made so that the repetition of the initial pattern will be repeated, and the wave pattern. Now let us introduce the rule of the wave excitation. Since each element has a high level of activity around it, we need to check whether there is a pattern of activity in his memory. If there is a pattern, then the element being excited or not will depend on the choices made in the beginning. The resulting logic of the automaton can be described as follows: if an element encounters an unknown signal, it either triggers or not on a random basis, and it stores the signal and the choice it has made; if the signal is known, it repeats its initial choice. The set of triggered elements generates the wave front, diverging from the pattern of the initial compact activity. The relaxation ensures unidirectional wave propagation from the places where activity occurred in the directions where it did not happen yet. This way we defined, constructed and have got an automaton that memorized a unique wave pattern, unambiguously corresponding to the original activity pattern. Repetition of the initial activity will not require elements to randomly determine their states. The elements "recognize" pattern they encountered before and propagate it further, thus, in the end, propagation of a wave happens with the same pattern as initially (Fig. 13.2b).When a different compact activity pattern appears, the automaton will generate the propagation front wave of activity exactly the same way. However, importantly, the new wave pattern will be unique and different from the previous wave pattern. Any compact combination of active elements will generate a unique wave pattern. For each emitting pattern, the propagating wave, firstly, will have a unique pattern different from all other wave patterns and, secondly, this pattern would always be the same for the same initial activity. It means that, if we define a glossary where each item is encoded by a compact pattern, then we will be able to transmit the information about activity of that item (i.e. its encoding pattern) across the plate of a cellular automaton. Indeed, as each initial pattern creates a unique wave pattern, it is possible to judge what notion the wave propagates in any arbitrary location of a cellular automaton plate. The Fig. 13.3 shows how patterns of wavefronts differ in the same location of a cellular automaton plate for two different initial patterns.

13.1.2 Properties of Information Waves

If a signal is encoded by a rather small percentage of activity across elements, then several waves can propagate through the automaton simultaneously without losing their individuality and not interfering with each other. During simultaneous spreading of several waves, the wavefronts of these waves can pass through each other keeping their pattern intact.

A binary vector can describe the activity of the automaton in each area. The crossing of signals forms a binary vector as the logical sum of each signal's binary vectors. It is equivalent to a Bloom filter (Bloom 1970). Accordingly, the false positive rate can be calculated the way it is for the Bloom filter. It is worth mentioning that the signals encoded by such an automaton gain a property of duality that corresponds to the wave-particle duality. Same as quantum-scale objects may be partly described in terms not only of particles, but also of waves, an informational signal in the simulation described acts as a pattern that triggers a wave, and a wave

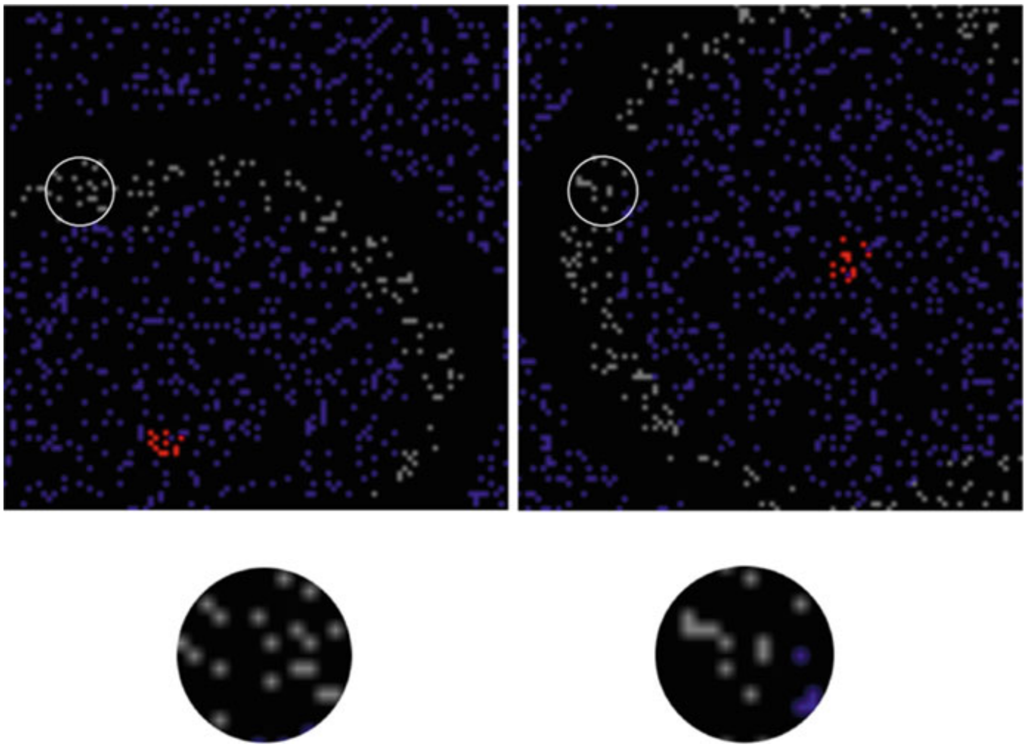

Fig. 13.3 The wave patterns from different initial patterns emerging at the same position of the cellular automaton

itself. In each phase of its propagation this wave forms a pattern which in turn propagates the wave further. The Huygens-Fresnel principle is applicable to the spreading of an information wave where each point reached by the wavefront could be considered as an independent source of emission of a spherical wave.

Let us pick an item known to the automaton. That item would match a wave with a unique pattern. If a fragment of this pattern is reproduced anywhere in the plane of the automaton, then this area will spread a wave reproducing the same unique pattern on all the way through its repetition. For example, if a specific pattern is created in the automaton in the area encircled by the line 1 (Fig. 13.4a), then the wave front will create a unique pattern for this wave reaching the area 2 (Fig. 13.4b). If a new wave is being emitted from the area 2 (Fig. 13.4c) by the pattern that was a part of the original wave then wave front reaching the area 1 re-creates the original pattern there (Fig. 13.4d).

This way the full connectivity of the automaton plate is achieved. Potentially, any area can store in memory any wave information and play it back later, all areas on the automaton plate reached by the corresponding wave, will have access to this information. The cellular automaton described here was modeled on a computer and showed stable operation over a wide range of parameters (Redozubov, Programs. [Online] http://www.aboutbrain.ru/programs/).

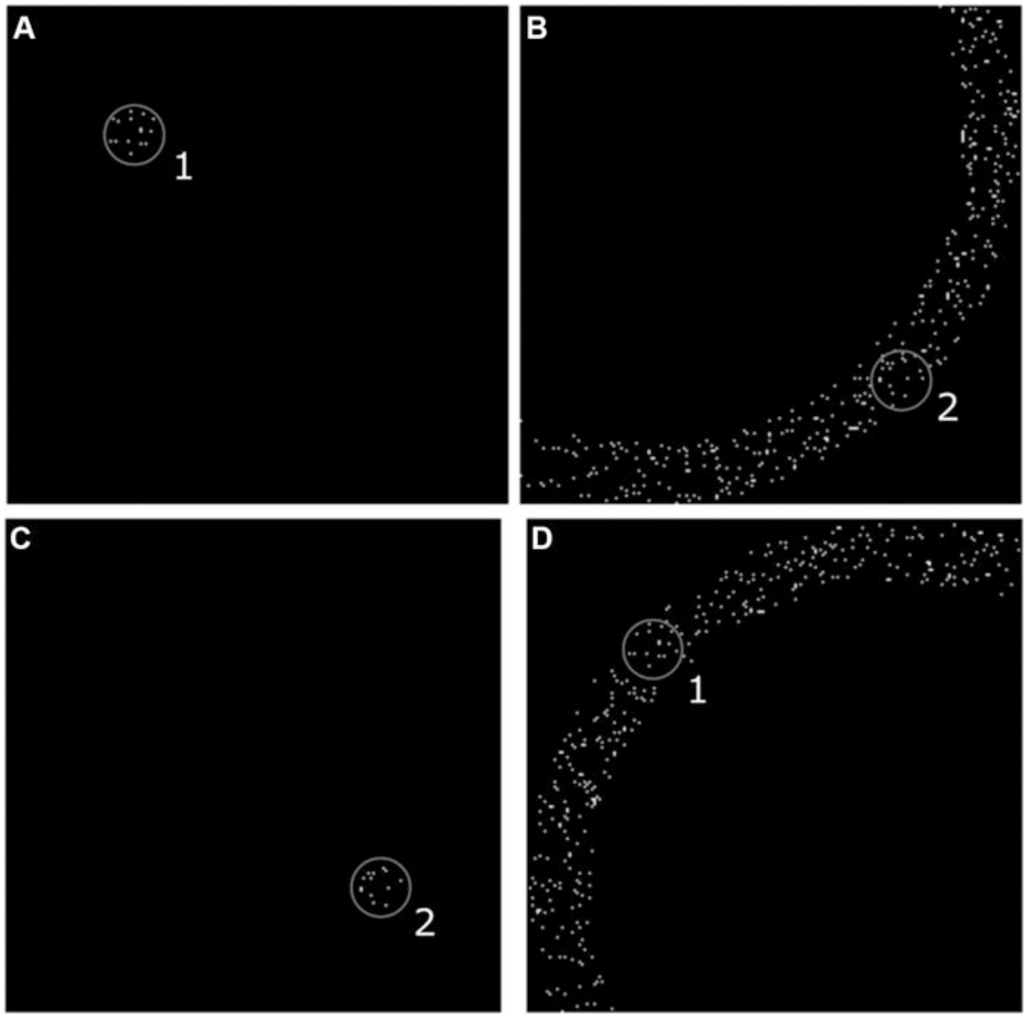

Fig. 13.4 A pattern emits a wave from the area 1 and, reaching the area 2, creates a unique pattern there (*top* figures). A new wave is being emitted from the area 2 by the pattern that was a part of the original wave. The wave front reaching the area 1 re-creates the original pattern there (*bottom* figures)

13.1.3 Brain Cortex Patterns

What structures in the cortex can act as elements of a cellular automaton? Such a structure should meet the following requirements:

- The structure can have at least two different states;
- There should be the capacity to transmit information about the state to the neighbors;
- There should be a mechanism to allow the structure under the influence of the pattern of neighbors to change its status;

- There should be a mechanism to selectively respond to different surrounding patterns;
- Transfer of information should be fast enough to match the rhythms of the brain;
- Since it is assumed that the pattern-wave mechanism should involve for every time transfer a large number of elements, the energy cost of each item should be minimal.

Thin branches of the dendritic trees are the most suitable candidate for these functions. According to neuron doctrine, the branches of the dendritic tree contribute to the functions performed by the neurons to which they belong. However, they can also have individual properties and act in some situations as autonomous elements. It has been shown that dendrites have cable properties (Wilfrid 1959). A branch of a dendrite can be compared with a cable, which has an internal resistance, leakage resistance and capacity of the surface. Although dendrite resistance very large and significant leakage, nevertheless currents that arise from excitatory postsynaptic potentials can have a significant impact on the overall state of the neuron. It can be assumed that the role of these currents is especially high at short distances, for example, within single dendritic branches of the tree.

A detailed mechanism of how the dendritic sections can act as carriers of information waves is described in (Redozubov 2016).

It can be assumed that the spread of dendritic activity patterns is accompanied by the appearance of spontaneous activity in certain neurons. This spontaneous activity can be interpreted as the calculation by local neuronal groups of the hash functions from the dendritic signals. Such spontaneous activity of neurons is very similar to the “neural avalanches” observed in the monkey cortex (Petermanna et al. 2009).

In the described model, all information arriving at any area of the cortex can be read by analyzing the state of any of it small fragments. This view on cortical processing suggests the possibility of brain-machine interfaces based on cortical microcircuits (Lebedev and Opris 2015).

13.2 Holography Memory

13.2.1 Patterns Interference in a Cellular Automaton

As described above, the cellular automaton wave activity, while spreading, creates a unique pattern on the automaton plate. A unique feature of our automaton is that reproducing a fragment of this wave anywhere in the automaton will cause the wave propagation from this point with exactly the same pattern as the original wave had. This means that the information encoded by a wave can be stored anywhere on the automaton plate by memorizing the pattern that occurs at that location when the information wave passes through. The elements of the described automaton have memory. The memory of an element is its ability to store certain patterns of activity within its field tracking area and then respond with its own activity whenever any of the stored patterns reappears.

The memory of elements can be used as a universal cellular automaton storage device that implements an associative array. The associative array is a storage of "key-value" pairs. To be able to manipulate the stored data, an associative array must support the following operations: to add a pair, to search for pairs (by key or data), and to remove the pair. For closer analogies to the cortex, let us turn from the flat cellular automata to the 3D by replacing the flat tracking field with volumetric. Let's place the elements in the nodes of a regular lattice. We assume that the automaton thickness is substantially smaller than its surface. Let's allocate for observations a cylindrical volume with dimensions comparable to the tracking field of the automaton elements (Fig. 13.5a). We call this the unit volume size, meaning that this is the minimum space that would guarantee a wave propagation, if a fragment of this wave is reproduced inside this volume.

Suppose that two information waves were sequentially emitted. The first wave carries a value that we want to store. The second wave is a unique key that will serve

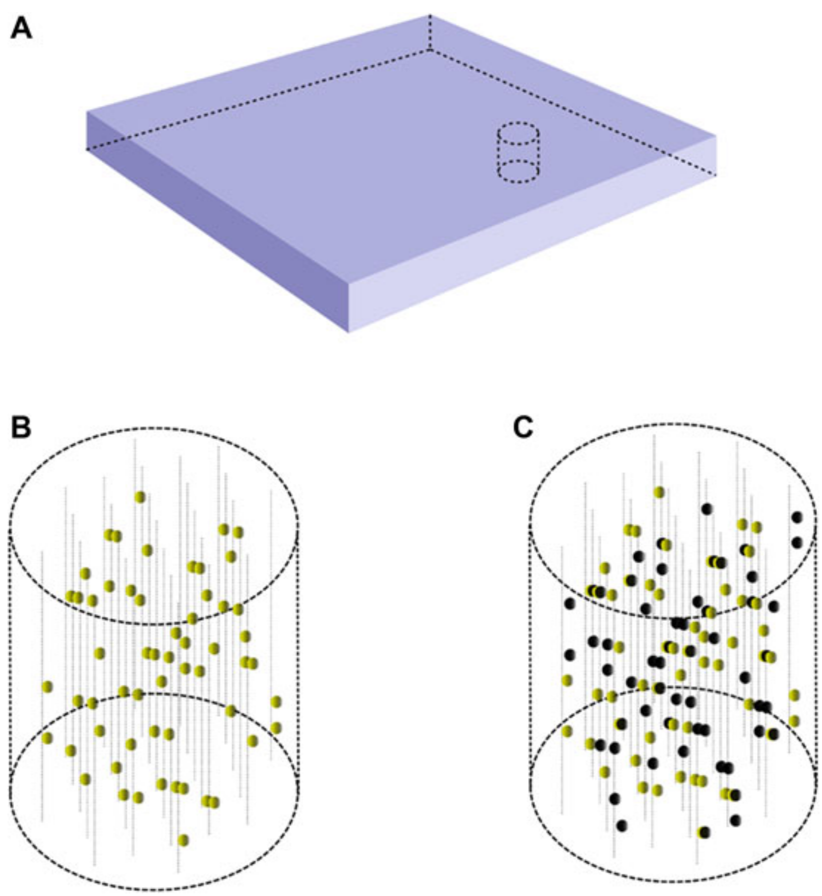

Fig. 13.5 (**a**) a spatial cellular automaton and the marked cylindrical fragment of a size comparable to the size of the tracking field of an element. (**b**) the trace of the information wave, carrying the value that will be remembered. (**c**) the trace of two waves. Elements that encode value are in *yellow*. Elements that encode key are painted *black*

as the identifier for the information stored. Each wave will propagate its pattern over the entire automaton space, that is, each area will contain two patterns formed by the first and the second waves, respectively. In the observed fragment, the first wave will leave a trace, as shown in the figure (Fig. 13.5b). The second wave will leave its trace in the same place. Let's mark elements of each waves with a different color, while some elements can receive two colors at once (Fig. 13.5c).

Now let's consider memorizations. For this purpose all yellow elements remember the pattern of the black elements. As a result of this kind of "interference", this area will memorize a "key-value" pair. Since we have chosen the volume comparable to the tracking area, the pattern enclosed therein can propagate its own wave if necessary. That is, if we subsequently reproduce the wave encoding key (the black wave), then the yellow items activate, because the pattern of black elements is the signal that causes their activity. As a result, a pattern encoding the value for the corresponding key arises in this volume. This pattern emits a wave that will spread the information retrieved from the memory space along the automaton. Actually, the described above is the implementation of store-and-search the information by a key.

If all keys are unique, the key propagation will cause a unique responding information wave corresponding to the paired value for that key. If the values are also unique, it is possible to do a reverse search of a key by its value. To store a single information element (a "key-value" pair), memorizing in a unit volume is enough. However, nothing prevents to store information "redundantly" distributed across the automaton. This means that the information is stored not in one place, but in the entire volume of a cellular automaton. Both local and distributed memory, as shown below, are extremely important for the implementation of information processes.

Interference of two waves of information and distributed storage makes the described mechanism extremely similar to optical holography (Gabor 1948). The main property of optical holograms is that each section of a hologram contains all the information about the entire light beam. The same property is incorporated in our memory model.

13.2.2 Special Dendrites Points

The pyramidal and stellate neurons constitute the main percentage of cortical neurons (Braitenberg and Schuz 1998). The axons of these neurons are characterized by highly branching collaterals. Most of the synaptic axon contacts falls on the volume of a size comparable to the size of the dendritic tree. This axon geometry ensures that the axon signal becomes available in almost all dendritic branches, located in a neighborhood (radius of the order of 50–70 microns) of the neuron.

The availability of a signal means that any dendritic branch in a proximity of a neuron has a segment close to the axon of this neuron. Accordingly, in a moment of activity of the neuron, when each synapse of its axon releases neurotransmitters, a portion of these neurotransmitters can reach the dendritic branch due to spillover.

The synapses surrounding a dendritic branch, both its own and external, are the sources of the extrasynaptical neurotransmitters for this branch.

In reference (Redozubov 2016) it is shown that with a probability close to 1 on any dendrite segment for each selected surround neurons signal will be a place in which will meet a minimum 5 of active axons of neurons. This place on the dendrite can be considered as the favorite in relation to the selected signal. To recall exactly which axons (synapses) had been active, it will subsequently with high accuracy detect repetition of the same signal.

13.2.3 Coding Signal in Selected Place by a Combination of Neurotransmitters

For the majority of synapses, at the time the activity is allocated a "basic" neurotransmitter, and in addition to one or more of the neuropeptide (Lundberg 1996; Bondy et al. 1989). The presence a large number of neurotransmitters and neuromodulators let us suggests, that the primary function of such a manifold – is the creation in time of synchronous neuronal activity in each point in space a unique combination of neurotransmitters and modulators. It can be assumed that the additional substances in synaptic vesicles distributed throughout the synapses so as to provide at each location a maximum space diversity among neighbors. If so, the detection of a particular combination of synaptic activity reduced to determining that the corresponding synapses unique set of emitted substances.

Thus, if in a location selected against a specific dendrite signal to place the detector, sensitive to the combination of substances, characteristic of this signal, the operation of the detector is very likely to represent the repetition of the original signal.

13.2.4 Receptors Neurons as Storage Elements

In addition to direct transfer mechanisms, there exist indirect mechanisms, which is activity-related metabotropic receptors. These receptors are not ion channels and, therefore, do not participate directly in the polarization or depolarization. The metabotropic receptors act indirectly by modifying the activity of ion channels, ion transporters and receptor proteins. Impact on the neuron's membrane potential metabotropic receptors is exerted through G-proteins (Dunlap et al. 1987).

Neighbor receptors can be connected, creating dimers. Dimers, in turn, unite to form clusters of receptors. The receptors cluster suitable for such role are neighbors synapses specific combination activity detectors. However, in order to use these receptors as universal memory elements, there must be mechanisms transform those receptors from sensitive state to insensitive state and back. Such mechanisms are detailed in (Radchenko 2007).

Therefore, there is a possibility to describe a hypothetical mechanism of memorizing that's based on patterns interference Assume in a local capacity of the crust we have two patterns of activity which sequentially replace by each other The first pattern describes an information, a second is an id. The both patterns consist of the great amounts of active dendritic elements. After performing hashing for the first pattern we've got a pattern of spikes synchronous activity of neurons. Active elements by the second pattern indicate dendritic segments, which have to memorize a pattern of neurons activity. On each segment, that has to memorize an image of volume activity, will found a favorites location relative to the signal. In favorite's location either through a random combinatorics already exists a ready cluster of receptors corresponding to the chemical composition by spillovers, or perhaps, such a cluster can be dynamically generated. The cluster's receptors are changing their conformation that brings the cluster into the sensitive state.

In this way, a pair of identifier pattern and its informational description in form of "hash" is memorized. Swapping of the information description and identifier, will result in memorization of the information description pattern in conjunction with the identifier "hash code".

Passage of such processes throughout the cortex will result in a "holographic" memory, where the same information is stored in each element.

It turns out that the clusters are receptive, sensitive to certain surround signals, potentially, may be memory elements that form a memory trace called engrams. Memory, created in such a way is time-dependent. The conformation of the receptor has hysteresis properties (Radchenko 2007). This means that receptors can stay in sensitive mode until external stimulus returns them to their original state. Such exposure may be, for example, a strong change in membrane potential. In this state, the memory status is short-term. That is the latest recording readily available for the memories, but the availability is reduced over time as going receptors reset. Engrams can be stored for long time. Adhesion and polymerization processes can fix receptor conformational changes. This is transforming memory into a state of sustainable storage. This memory can be stored until the end of life.

In a spatial structure that interlaces axons and dendrites, and which employs the principle of "favorite locations", the memory elements could include various types of receptors. This means that, most likely, the majority of membrane receptors are associated with any working memory systems. Furthermore, glial cells of the cortex have the same sets of receptors as neurons (Halassa et al. 2007), and thus can participate in mechanisms of memory. Astrocytes are able to both enhance the reaction of the synapse due to release of the corresponding mediator, and to weaken it by its absorption or release of the neurotransmitter binding proteins. In addition, astrocytes are capable of releasing signaling molecules that regulate the release of the neurotransmitter axon. Concept signaling between neurons that takes into account the effect of astrocytes, is called tripartite synapse (Fields and Stevens-Graham 2002). It is possible that the tripartite synapse is the main element that implements the mechanisms of mutual work of the various memory systems.

13.2.5 The Hippocampus Role. IDs for Information

In 1953, bilateral hippocampus resection was made as anti-epilepsy therapy for patient that is known as H.M. (Henry Molaison) (Scoviille and Milner 1957). As a result, H.M. had lost ability to memorize anything. He remembered all that were happened to him before the operation. However, new memories became completely lost on his attention switch. H.M. case is unique. In other cases of hippocampus resection, without full both sided destruction memory corruption were not so well presented of were not existed at all (Scoviille and Milner 1957). Full hippocampus resection makes forming new memories impossible. Hippocampus dysfunction may lead to Korsakov syndrome, which appears as impossibility to fix current events while old memory is safe.

The widely known hippocampus role is holding current memories and reorganization later within cortex space of this memories. In the describing model hippocampus have different role. It is memories unique keys creation. The keys created by the hippocampus are distributed to the corresponding cortex zones through projection system. Interference of hippocampus identifiers and information descriptions creates the memory. Thus, memory forms "in its own place" and does not move between the hippocampus and cortex. Such representation agrees with the experimental data quite well. Hippocampus removal makes new memories formation impossible because of memories keys lack. Old memories stays untouched as they are independent of hippocampus. Their identifiers may be extracted and used without hippocampus action.

But the main arguments in favor of described hippocampus role connected with functions found in hippocampus and have no direct relationship with memory mechanism. In 1971 John O'Keefe discovered place cells in hippocampus (O'Keefe and Dostrovsky 1971). These cells act as inner navigation system. If a rat is placed in long hall, then it is possible to determine rat particular place from particular cells activity. What is more that cells activity is independent of way rat is come in particular place.

Hippocampal formation contains neurons that encode spatial location using grid-like coordinates (Hafting et al. 2005). In 2011, it was revealed that there are cells in hippocampus that code time intervals in the same way. Their activity forms rhythmic patterns even if there is nothing happening around (MacDonald et al. 2011).

Storing data in form of key-value pairs creates associative array. In an associative array, key has two functions. It is unique identifier, which lets differ one key-value pair from another. In other side key may hold information that can make search simpler. As example, PC file system may be considered as associative array. Value is information in the file; key is information about file. Information about the file is the path that defines storage place, name of the file, date of creation. For photos additional information – geotags, place where picture was taken, may exists. For music files there are album name and performer name. All this data about files forms complex keys that is identifies file uniquely but at the same time lets perform searching by any key field or its combination. The more key details, the more flexible search ability is.

As brain implements the same informational tasks as computer systems, it is reasonable to make an assumption that storing data in key-value pairs by a brain will lead to creation of the keys which will be more convenient for searching.

For human memories it is reasonable to have the following key descriptors:

- Scene designation;
- Position in space designation
- Time designation
- Number of concepts that is related to what is happening. Something like article keywords which describes article content.

It looks like the hippocampus not just works with the scene, position in space and time, but uses it for composing complex informational keys for the memories. At least this explains why such different functions came together in one place that is responsible for memory formation.

Time encoding is of special interest. Human memory let remember not only static images, but a sequence of scenes with their chronology. So, memory coding system must contain such ability. It was shown that the hippocampus has time cells that creates rhythmic patterns (MacDonald et al. 2011). Patterns cyclicity suggests that hippocampus may use the same principles for creation identifiers time fields as humans do for time measuring.

13.3 Algorithmic Model Based on the Meaning of Information

13.3.1 Cryptography and Meaning

Consider an example from the field of cryptography. Suppose that we have a stream of encrypted messages. The encryption algorithm is based on substituting characters of the original message with another according to rules, which are defined by encrypting mechanism and key. With something like that has dealt Alan Turing, hacking the code of the German “Enigma”. Suppose that there is a finite set of keys and that we know the algorithm of the encrypting mechanism. Then, to decrypt the current message you need to iterate over all the keys, decode messages and try to find meaningful among them.

To determine the meaning of the message, it is necessary to have a dictionary with words that may appear in the message. As soon as the message will take the form in which the message words coincide with the words from the dictionary, it will be possible to say that we found the right key and decrypted received message. If we want to speed up key search, then we will have to parallelize the decoding process. Ideally, you can take as much parallel processors as number of different keys. Allocate keys on the processors and run on each transformation reverse conversion with his key. Then check the result for the meaning. In one passage of the calculations, we will be able to check all the possible hypotheses about the key used to find out which one is most suitable to decrypt the message.

To test the meaningfulness each of the processors should have access to a dictionary of possible words in the message. Another option – each processor must have a copy of the dictionary and turn to it for checking. Let us consider the second option. Now, make the task more interesting. Suppose that we know only a few words for meaningfulness test, which constitute our dictionary. Then in the message stream, we can find the key only to those in which there is at least one of the known words. There may be situations when multiple keys will show words in the decoded message, which we have in the dictionary. Then, you can either ignore such messages as undeciphered, or select the key that gives a greater match of words in the dictionary. When we find out the right code for these few messages, we will get the correct spelling for the other previously unknown to us words. These words can supplement vocabulary of the processor, which had found the right answer. Furthermore, new words can be transferred to all other processors in addition to their local dictionaries. As you gain experience, we will decrypt greater percentage of messages until we get a complete dictionary and close to one hundred percent decryption effectiveness.

The resulting cryptographic system is interesting because it allows us to introduce the notion of "meaning" and give an algorithm that allows to work with it. Sense for such system is a property of the encoded message that appears in the selection of such a code, which creates a decoded message, interpreted on the existing dictionary. For described cryptographic task the meaning of the encoded message can be called the couple "key-decrypted content." A proper understanding of the meaning of the message – is the selection of the same code, and obtaining the same messages that were laid by the sender. The algorithm for determining the meaning – is to check all the possible interpretations and the selection of one that looks the most plausible in terms of memory, which stores all previous experience of interpretations.

13.3.2 The Meaning of Discrete Information. Frames

Interpretation of the meaning and the algorithm of its determination imposed for cryptographic tasks can be extended to the more general case of arbitrary information messages composed of discrete elements. We introduce the term "concept" – c (concept). We assume that we have N available concepts. A set of all available concepts forms a dictionary.

$$C = \{c_1 \cdots c_N\}$$

We define information message as set of concepts length k

$$I = (i_1 \cdots i_k)\text{ , where } i_j \in C$$

We assume that the message can be associated with his treatment I_{int} (interpretation). The interpretation of messages – it is also an informational message, consisting of concepts from set C. We introduce the rule for interpretations producing. We believe that any interpretation is obtained by replacing each concept of the original message with some other concept or with itself. Assume that exists a system for performed replacements, which is generally not known to us.

Let us introduce the notion of "subject" S. We define subject's memory as an array of information known to him, received interpretation. Information with the interpretation can be written as a couple

$$m = \left(I, I^{int}\right)$$

Then, the memory can be represented as:

$$M = \{m_i |\, i = 1 \cdots N_M\}$$

Determine the first stage of subject's learning. Perform supervised learning. We submit informational messages and their correct interpretation. Memorize all the information received. Based on the memory, formed supervised, we can try to find a system in comparison concepts and their interpretations.

Firstly, we can draw up a range of possible interpretations for each concept. To do this, for each concept we need to collect all its interpretations that are stored in memory. By the way, the frequency of using a particular interpretation can give the appropriate estimate of the interpretation probability. Secondly, we can use any reasonable method to solve the problem of clustering and divide m_i objects into classes, according to how the same concepts are interpreted in the class. We will try to make sure that all objects within the class use the same interpretation rules for the constituent concepts. Let us call classes resulting from a clustering – "contexts".

The set of all contexts for the subject S forms the space of contexts $\{Cont_i\}$. For each context i, you can specify a set of rules for interpretation of concepts

$$R_i = \left\{ \left(c^{orig}, c^{int}\right)_j \middle| j = 1 \cdots N_{Context} \right\}$$

After finishing supervised learning, we can introduce the new algorithm that allows us to interpret new information. We distinguish memory M^{int} from M, consisting solely of interpretations

$$M^{int} = \left\{ I_i^{int} \middle| j = 1 \cdots N_M \right\}$$

We introduce a measure of coherence of interpretation and memory of interpretations. In the simplest case, this may be the number of matches interpretation and memory elements, i.e. the number of times it occurs in such an interpretation in interpretations memory

$$\rho(I) = \sum_i \begin{cases} 1, I = I_i^{int} \\ 0, I \neq I_i^{int} \end{cases}$$

Now for any new information I for each context $Cont_j$ we can get an interpretation I_j^{int}, applying to the original data transformation rules R_j. For each of the resulting interpretations can determine its consistency with interpretations memory

$$\rho_j = \rho\left(I_j^{int}\right)$$

The scheme for the calculations of the context is shown in Fig. 13.6a.

We introduce the probability of interpretation the information in the context j.

$$p_j = \begin{cases} 0, \rho_j = 0 \\ \rho_j / \sum_i \rho_i, \rho_j \neq 0 \end{cases}$$

As a result, we will get the interpretation of the information I in each of the K possible contexts and the probability of this interpretation

$$\left(\left(I_1^{int}, p_1\right) \cdots \left(I_K^{int}, p_K\right)\right)$$

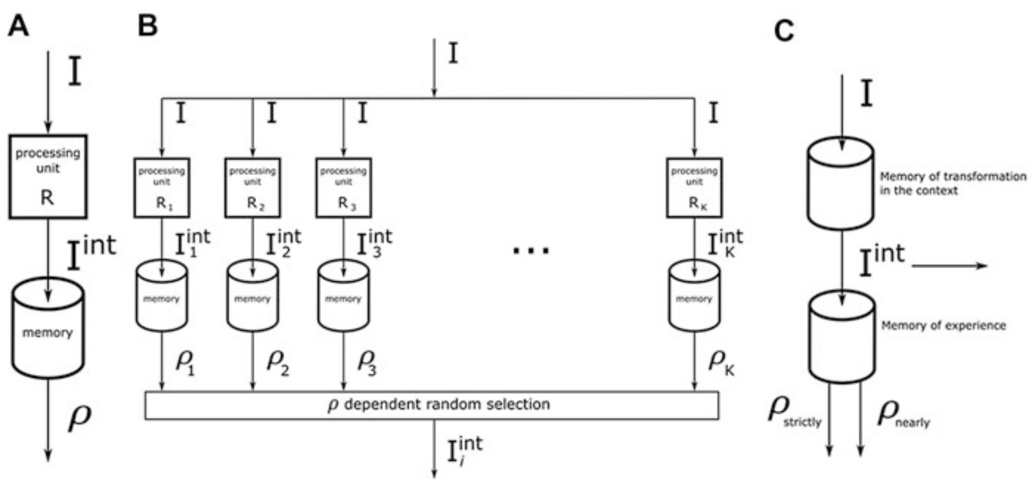

Fig. 13.6 (**a**) computational scheme of the context module. (**b**) computational scheme for determining one of the meanings in the system with K contexts. (**c**) diagram of basic computing functions for minicolumn of cortex

If the probability is zero, we can state that the information is not understood by the subject and has no meaning for him. If there is a probability different from zero, then the corresponding interpretation form a set of possible meanings of the information. If we decide to determine the one main information interpretation for the subject, you can use context with the maximum probability value. As a result, the notion of "meaning" can be described as follows. The meaning of the information I to the subject S is a set of interpretations that the subject finds during context matching, which was built on the determination of the conformity of interpretations that have arisen in different contexts and memory. General computational diagram associated with the meaning determining can be represented as a set of parallel working contextual computing modules (Fig. 13.6b). Each module performs the interpretation of the original description by its own transformation rules. The memory of all the modules has identical content. Comparison with memory provides the conformity assessment of the interpretation and experience. Meaning selection is based on the probabilities of interpretations. The procedure is repeated several times to retrieve the set of possible interpretations.

Once the meaning of the information is defined, memory can be supplemented by new experience. This new experience can be used for determining the interpretation of subsequent information and to clarify the context of space and transformation rules. Thus, you should not allocate a separate stage of primary education, but simply accumulate experience, while improving the ability to retrieve the meaning.

Described semantic approach to data contains a few key points:

- Information descriptions to which this approach is applicable are built from discrete (nominal) terms. This is determined by the ideology of the concepts comparison and their interpretations in a certain context. Descriptions of different nature, for example, quantitative indicative descriptions can be used only after conversion of the quantitative variables to their approximate discrete representation.
- Experience allows you to create the space of contexts and interpretation rules in these contexts. Accordingly, the meaning can be determined by the subject only with a certain experience.
- Since experience of different subjects can vary and different meanings can be obtained as a result of perception of the same information;
- Information can be specially prepared by the sender so as to maximize the probability of the specific meaning for recipient;
- Needless to say that the information contains meaning regardless of the perceiving subject.

Meaning is the result of "measurement" of the information made by the subject. Prior to determining the meaning, for a specific subject information contains interpretations in all contexts for which was a non-zero probability of these interpretations. Each "measurement" allows you to see one of the possible meanings.

Described context meaning model, in many respects, solves the same problem as the concept of Marvin Minsky's frames (Minsky 1974). Common challenges, that models are facing, will inevitably lead to similar realizations. Describing the frames, Minsky uses the term "microworlds", meaning by them situations in which there is a certain consistency of definitions, rules and actions. These microworlds can be compared with the contexts in our definition. Selection of the most successful frame from the memory, and adaptation to the real situation, the same can be largely matched to the procedure of determining the meaning.

When using frames to describe visual scenes frames are treated as different "points of view". In this case, different frames have a common terminal that allows you to coordinate information between frames. This corresponds to the way, how in different contexts interpretation rules may lead different initial description to the same descriptions-interpretations.

Popular in the programming, object-oriented approach is directly related to the theory of frames. It uses the idea of polymorphism, when the same interface when applied to objects of different types causes different actions. It is close enough to the idea of interpretation the information in context.

Despite the similarity of approaches associated with the need to answer the same questions, context-semantic mechanism differs significantly from the theory of the frames, and, as will be seen below, cannot be reduced to it.

The main value of the context-semantic approach is that it is equally well applicable to all kinds of information faced by the brain and operates. All the zones of the real cerebral cortex extremely similar in terms of the internal organization. It makes one think that they all use the same principle of information processing. It can be assumed that the semantic context approach is that general principle.

13.3.3 Semantic Information

Words that build speech can be interpreted in different ways depending on the overall context. However, for every word is possible to make a range of values described in other words. Consistent variation of interpretation of the words allows to allocate the available set of contexts. Contexts can be associated with time, number, gender, subject area, topics and so on. Determining the meaning of the phrase – is the choice of the context and interpretation that the most plausible, based on the experience of one who tries to understand the meaning. There may be a situation where the same phrase in different contexts create different interpretations, but these interpretations will be allowed on the basis of previous experience. If the task is to determine the only meaning of this phrase, it is possible to choose the interpretation that has higher match to memory and, respectively, the calculated probability for it. If the phrase was originally formulated as ambiguous, it is appropriate to accept each of the senses individually and establish the fact that the author of the phrase intentionally or unintentionally managed to combine them in a single statement.

Natural language is a powerful tool for expression and conveying meaning. However, this power is achieved due to ambiguities and interpretations of probabilistic nature, depending on the experience of the perceiver. When the meaning of the phrase is rather complicated, as it happens often enough, for example, when discussing the scientific or legal matters makes sense to switch to the use of special terms. Transition to terminology is the choice of the interlocutors a coordinated context in which the terms are treated equally by interlocutors. For this context to be available to both interlocutors each of them needs some relevant experience. In order to interpretations were similar its require a certain similarity of experience (learning).

For the natural language, it is possible to use a measure of consistency of the context and memory based not only on complete coincidence of descriptions, but also on their similarity. Then becomes available a larger number of possible meanings and there is a possibility of additional interpretations of phrases. For example, so you can correctly interpret phrases containing errors or internal contradictions. In determining the meaning is easy to take into account the overall context. For example, if the phrase allows interpretation in different contexts, preference is given to contexts that were active in the previous sentences and established the general context. If the phrase allows interpretation only in the context different from the main, it will be taken as shifting to another topic.

13.3.4 Auditory and Visual Information

Analog audio signal can be easily converted into a discrete form. First the discretization of time when a continuous signal is replaced by measurements performed with a sampling frequency. Then, quantization of the amplitude is done. Wherein signal level is replaced by the number of the nearest quantization level. The resulting recorded signal can be divided into time intervals and for each applied windowed Fourier transform. The result is a sound encoded as a series of spectral measurements.

Let's define a cyclic identifier with period N_T for time intervals The first N_T intervals will then be numbered from 1 to N_T. The N_{T+1} interval will then be numbered 1 again, and so on. Thus, interval numbers will be repeating every N_T intervals. Let's assume that a Fourier transform contains N_F frequency intervals. Then, any spectral measurement will consist of N_F complex values. Let's replace every complex value with its respective amplitude and phase and apply quantization. Number of quantization levels is N_A for amplitude and N_P for phase. Within the range of N_T time intervals each spectral record element may be described with the following set: (time interval code, frequency value, amplitude value, phase value).

Let's introduce the set of concepts C, allowing us to describe the sound within the period of the cyclic identifier. This set will include all the possible combinations of type

$$(n_T, n_F, n_A, n_P)$$

Total number of such concepts will

$$N = N_T N_F N_A N_P$$

The set of concepts C will then contain N elements

$$C = \{c_1 \cdots c_N\}$$

Any sound signal not longer than N_T time intervals can thus be recorded as the sequence of concepts.

$$I = (i_1 \cdots i_k)\ \text{, where } i_j \in C$$

In practical applications, e.g. in speech recognition, the signal may be transformed according to some set of rules, while still retaining its original meaning, i.e. the words it represents.

The simplest signal transformations are:

- Time shift
- Frequency shift
- General volume change
- Time scale change (reproduction speed change)

Let's introduce the context space, covering all possible transformation combinations. For each context transition rules can be defined, i.e. we can describe how each of the source concepts will look like in the context of the appropriate transformation. E.g. for the context, shifting the frequency one position up all the concepts will get interpretation, shifting their frequency one position down. Pure 1 kHz tone is equal to 900 Hz tone in the context of common frequency shift 100 Hz up. The same applies to other transformation types. After the transition rules for different contexts are described it becomes possible to recognize words regardless of how they are transformed. The moment of pronunciation, loudness, voice pitch and speed will not affect the possibility to compare current information with the memorized one. Current sounding will be transformed to different interpretations, corresponding to all possible contexts. In the context, corresponding to the appropriate transformation, description will result in an interpretation that could be easily recognized as being heard earlier.

In practice, while working with complex signals, like speech, single processing step is not sufficient. Initially, it is useful to separate simple phonemes. Contexts will then contain rules for transformations of simple sounds, as shown above. Then we can compose a description, consisting of phonemes. Here phonemes are complex elements, identifying not only the sound form, but also it's pitch, timing and pronunciation speed. Information consisting of phonemes can then be further processed on the space of own contexts and own memory. Sophisticated contexts may be constructed, not just limited by simple transformations. The definition of appropriate context in itself creates additional information. For example, in case of speech different intonations and language accents are contexts. Not only intonation and accent contexts increase the precision of speech recognition, but also give additional information on how the phrase was told.

Similar considerations are valid for visual information (Redozubov 2016). The visual description may correspond to a certain binary code. Different image transformations are the rules for changing this code. A set of various transformations, for example, horizontal and vertical displacements, and rotations creates a space of visual contexts.

The basic idea of this approach lies in the fact that for the invariant representation of an object is not necessarily to spend a lot of time for training, showing the object from different angles. It's much more efficient to teach the system the basic rules of geometric transformations inherent in this world and common to all objects. Partially the described approach is implemented in a well-proven convolutional network (Fukushima 1980; LeCun and Bengio 1995).

13.4 Cortical Minicolumns

13.4.1 The Memory Capacity of One Minicolumn

The cerebral cortex is composed of minicolumns. Minicolumn is a vertically spaced group of 80–120 neurons.

Previously it was shown the formation of a memory circuit, built on the interference of two wave patterns. The first pattern defines the elements (dendritic section), which should keep the memory. The second pattern defines a memory key. Hash conversion from the second pattern creates a short key of memories (spiking activity of neurons). In special places receptor clusters that are specific to the combination of neurotransmitters has arisen, fixed memory. In (Redozubov 2016) show that with this approach, one mini-column of the cortex can store about 300 MB of information.

The approach based on the plasticity of synapses provides a much more modest result as the main memory element. The minicolumn contains about 800,000 synapses. Even assuming that the synapse due to changes in the level of plasticity encode multiple bits of information, obtained value will total only the hundreds of kilobytes. Increasing memory capacity three orders of magnitude gives a qualitative leap in information capabilities minicolumns. Since the nature of information stored in minicolumn, is close to the semantic information, the 300 MB capacity are quite sufficient to save, for example, all human memories that accumulate in the course of life.

The book of 500 pages in the uncompressed form is about 500 KB. A one minicolumn allows to store memories library consisting of 600 volumes. Approximately one book per month of life, or 15 pages per day. It seems that it is enough to hold the semantic description of everything that happens to us.

Three hundred megabytes of memory minicolumns should not be compared with the gigabyte range photographic libraries or film libraries. When the image stored in the memory, it not stored in photographic form. It can be stored in the form of short semantic description consisting of concepts corresponding to the image. At moment of memories image is not reproduced, it reconstructed anew, creating the illusion of photographic memory. This can be compared with the way the human portrait can be restored close enough to photograph only according to his verbal description.

At the first moment the idea that only 100 minicolumns neurons can store the memories of a lifetime seems absurd, especially for those who are accustomed to believe that memory is distributed over the entire space of the cortex. Moreover, the duplication on many millions of minicolumns of the same information in the traditional approach seems pointless waste of resources. However, the meaning approach allows us to take under such an architecture of cortex serious justification.

13.4.2 The Basic Computing Functions of Cortical Minicolumns

The basic idea of defining the operation of a minicolumn, is quite simple (Fig. 13.6c). Consider one operating cycle of cortex. The information which carrying the current description, is distributed by zone of cortex consisting of a plurality of minicolumns. Each minicolumn sees this information as patterns passing therethrough defined activity (presumably activity of dendritic segments). Each minicolumn keeps the memory of transformations. Each of the minicolumn is responsible for its own perception of the information context. Each context implies its own, distinct from others, rules of transformation source descriptions in their interpretation. Minicolumns memory of transformations – a mechanism for the transformation of patterns of basic concepts, that constitute the description, in the patterns of concepts relevant in the context of the interpretation of a particular minicolumn.

As a result of the transformation of concepts constituting present description, in their context-dependent interpretation in each minicolumn appears own hypothesis of interpretation of this description. This hypothesis is a description, composed of corresponding interpretations. Physically, it most likely, looks like accumulated over clock cycle the activity of dendritic segments. This activity can be associated with a binary array, composed on the basis of Bloom filter. It can be assumed that there are mechanisms that allow the transformed information and the original description co-exist without interfering with each other. It is possible that for a separate processing correspond to different layers of the cortex. The combination of the activity of dendritic segments leads to spike activity of minicolumns neurons. The code compiled from the activity of neurons, can be interpreted as a hash function of information description, corresponding to the interpretation of the original information.

Previously, it has been shown that the combination of neuronal activity may be the key to which the memory can be retrieved previous experience with the same or such like him the key. Memory of each minicolumn stores all the events previously. Previous experience, supposedly, stored as pairs "hash of information description – the identifier" and pair "hash of identifier – information description" Cloning the same memory on all minicolumns necessary so, that each minicolumn could compare own interpretation of information with all previous experiences.

It can be assumed that the result of comparing the current interpretation of the description and the memory is the calculation of the compliance functions. The first function of compliance indicates the presence of an exact match interpretation and some memory elements. The second function evaluates the overall similarity of interpretation and experience stored in the memory.

Signals compliance functions potentially can be encoded changes in the membrane potential of individual neurons or groups of neurons.

Matching functions allow to judge about how minicolumns context appropriate for the interpretation of the current information, that is, how much interpretation received in this context, in line with earlier experience. Making a comparison between minicolumns, you can select a minicolumn-winner.

The interpretation received at a winning minicolumn is based on the concepts common to the entire cortex. The winning treatment can be pattern-wave method distributed throughout cortex area and memorized by all cortical minicolumns.

The above description of the interference information and the identifier will allow to fill up the memory of each minicolumns "correct" interpretation of the new experience. New information will be comparing with that interpretation.

Winning minicolumns elements can, depending on what is required, to reproduce the relevant interpretation or current information or the most appropriate under the current description memory of past experience or even any information stored in the minicolumn. Reproduced information can spread on cortex area or can be projected to other areas for further processing.

If the initial information allows for multiple interpretations, all of them can be produced in succession. To do this, after the determination of the first interpretation suppress the activity of the relevant context and repeat the context selection procedure. So one by one, you can single out all the possible semantic interpretation of the analyzed information.

A cortical minicolumn in our approach is a universal module that performs and the autonomic computing, and interaction with others minicolumns. However, different areas face different information tasks. In some tasks more important is number of contexts and less important volume of internal memory of minicolumn.

In others, conversely, more important is volume of internal memory and as a consequence an increase of internal bit hash code, i.e., the number of neurons in the minicolumn. The optimal setting of universal computing modules for task specific areas of the cortex can go two ways. Firstly, the number of neurons in the minicolumn may vary for different areas of the cortex. Second, the potentially possible to combine several vertical columns of neurons in one computation module. The scope of axonal and dendritic trees of neurons, constituting a diameter on the order of 150 microns, can combine multiple columns into a single computer system without altering the above general principles of work.

In addition, it can be assumed that a full copy of the memory cannot fit into one minicolumn and be distributed in the space of a few neighboring minicolumns. Since dendritic tree diameter is about 300 microns, this space is potentially available for minicolumn for operation with memory.

Acknowledgments The author expresses his deep gratitude to Ioan Opris, Dmitry Shabanov and Mikhail Lebedev for the constructive discussion and assistance in the preparation and translation of this article.

References

Bloom BH (1970) Space/time trade-offs in hash coding with allowable errors. Commun ACM T 13(7):422–426

Bondy CA, Whitnall MH, Brady LS, Gainer H (1989) Coexisting peptides in hypothalamic neuroendocrine systems: some functional implications. Cell Mol Neurobiol 9:427–446

Braitenberg V, Schuz A (1998) Cortex: statistics and geometry of neuronal connectivity, 2nd edn. Springer, Berlin

Dunlap K, Holz GG, Rane SG (1987) G proteins as regulators of ion channel function. Trends Neurosci 10:244–247

Fields RD, Stevens-Graham B (2002) New insights into neuron-glia communication. Science 298:556–562

Fukushima K (1980) Neocognitron: a self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol Cybern 36(4):193–202

Gabor D (1948) A new microscopic principle. Nature 161:777–778

Gardner M (1970) Mathematical games – the fantastic combinations of John Conway's new solitaire game "life". Sci Am 223:120–123

Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. Nature 436:801–806

Halassa MM, Fellin T, Takano H, Dong J-H, Haydon PG (2007) Synaptic islands defined by the territory of a single astrocyte. J Neurosci 27:6473–6477

Lebedev M, Opris I (2015) Brain-machine interfaces: from macro- to microcircuits. In: Recent advances on the modular organization of the cortex. Springer, Dordrecht

LeCun Y, Bengio Y (1995) Convolutional networks for images, speech, and time-series. MIT Press, Cambridge

Lundberg JM (1996) Pharmacology of cotransmission in the autonomic nervous system: integrative aspects on amines, neuropeptides, adenosine triphosphate, amino acids and nitric oxide. Pharmacol Rev 48:113–178

MacDonald CJ, Lepage KQ, Eden UT, Eichenbaum H (2011) Hippocampal "time cells" bridge the gap in memory for discontiguous events. Neuron 71:737–749

Minsky M (1974) A framework for representing knowledge, MIT-AI Laboratory Memo 306. Massachusetts Institute of Technology A.I. Laboratory, Cambridge

O'Keefe J, Dostrovsky J (1971) The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. Brain Res 34:171–175

Petermanna T, Thiagarajana TC, Lebedevb MA, Nicolelisb MAL, Chialvoc DR, Plenz D (2009) Spontaneous cortical activity in awake monkeys composed of neuronal avalanches. Proc Nat Acad Sci 106:37

Radchenko AN (2007) Information mechanisms of the brain. St. Petersburg: s.n

Redozubov A (2016) The logic of consciousness. [Online]. https://habrahabr.ru/post/308268/

Redozubov A. Programs. [Online] http://www.aboutbrain.ru/programs/

Scoviille W, Milner B (1957) Loss of recent memory after bilateral hippocampal lesions. J Neurol Neurosurg Psychiatry 20:1

Von Neumann J, Burks AW (1966) Theory of self-reproducing automata. University of Illinois Press, Urbana

Wilfrid R (1959) Branching dendritic trees and motoneuron membrane resistivity. Exp Neurol 1:491–527