Physics – Uspekhi **61** (9) 871 – 895 (2018)

PHYSICS OF OUR DAYS PACS numbers: 87.19.L –, 87.85.St, 89.20.Ff

The robot and the human. Where's their similarity limit?

G R Ivanitskii

DOI: https://doi.org/10.3367/UFNe.2018.03.038302

Contents

Abstract. It is shown that the primary goal of robot–human collaboration is to ensure dynamic stability under varying environmental conditions. Compared with the robotic computer 'brain', its human counterpart has a multilevel hierarchical organization, with information processing occurring at all levels — from the quantum up to the social. Humans themselves set goals and improve the virtual model synthesized by their brains. The human brain can work by simultaneously using both classical deterministic logic and dialectical probabilistic logic.

Keywords: android robot, adaptation, creativity, memory, hierarchy of organization, classical and probabilistic logic

1. Introduction

1.1 Ambivalent attitude toward robotics

We are witnesses of exponential progress in robotics engineering. This article explores the comparison between the *artificial intelligence* (AI) *of creative android robots*[1] (CARs) and *human intelligence and how far the similarities go*. Developers of CARs hope that machines of this class will mark the greatest achievement of humankind in science and technology. But society adopts a mixed attitude toward this work.

On the one hand, intelligent robots are expected to be capable of doing a variety of work much better, faster, and at a lower cost than humans. This may result in workplace reduction and even a so change in the social structure of human society that its stability will break. Moreover, defense and law enforcement authorities may use CARs to achieve their specific purposes by force interventions. Ideas of hybrid systems integrating the human brain and AI are becoming increasingly popular. For example, Neuralink Corp. is developing brain–computer interfaces. Such systems are likely to bring forth a caste of 'superpolicemen' and 'supersoldiers'. The main risks associated with the implementation of these projects are potential faults and failures of hardware and software, which cannot be ruled out even though modern software programs and computers usually operate more reliably than humans, and when duplicating do not make mistakes. Nevertheless, they are vulnerable to hacker attacks and other chance interferences or intentional misadjustment. The high operation speed of CARs does not leave time for an

G R Ivanitskii Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences,
ul. Institutskaya 3, 142290 Pushchino, Moscow region,
Russian Federation
E-mail: ivanitsky@iteb.ru

Received 17 October 2017, revised 14 February 2018
Uspekhi Fizicheskikh Nauk **188** (9) 965 – 991 (2018)
DOI: https://doi.org/10.3367/UFNr.2018.03.038302
Translated by Yu V Morozov; edited by A Radzig

[1] *Creativity* — the ability to produce original and unusual ideas or to make something new or imaginative (*Cambridge English Dictionary*). *Android* (Greek ἀνήρ — *human* + suffix-old meaning 'resembling', 'like' = *humanlike*) — a humanoid robot or synthetic organism designed to act like a human.

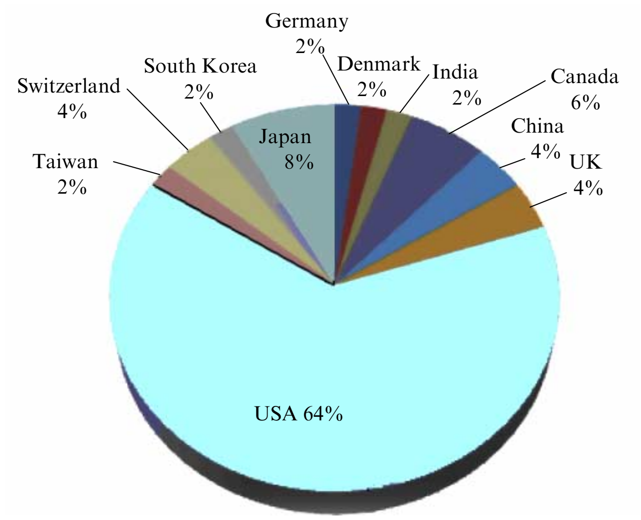

Figure 1. Diagram illustrating activity of 50 robotics companies in ten countries as of 2015 (after Editors of the site roboticsbusinessreview.com).

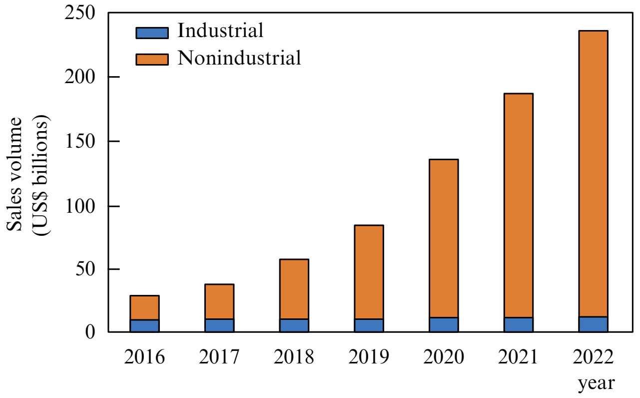

Figure 2. Expected growth in robot sales by 2022 is estimated at $237.3 billion. The 5-fold growth is mainly due to purchases of nonindustrial devices [Tractica: Total Industrial and Non-Industrial Robotics Revenue, World Markets (2016–2022)].

operator to manage the consequences of a failure, which may have a disastrous outcome. The sets of dangers associated with the development of creative AI are frequently likened to Pandora's box.

On the other hand, the future is created today. There is reason to believe that CAR programs will interact with us without malicious intent, being under strict human control, and the robot–human collaboration will bring great benefits to humankind. There is no shortage of predictions about how artificial intelligence is going to reshape where and how people work in, say, 2040, to say nothing of 2100. It is impossible to predict the relationships between people and new generations of supercomputers implementing AI and CARs, nor can humans' voracious curiosity be limited and the scientific advancement it promotes stopped. Attempts to prohibit research in certain fields made in the past (e.g., artificial fertilization of a human ovum or gene manipulation) slowed the development of science but could not arrest it [1]. Robotics engineering is gaining momentum in all developed countries (Fig. 1).

The process is so rapid that this diagram looks obsolete. China has made considerable progress in this field in the last 3–4 years. The majority of technical universities in Russia are involved in a national robotics engineering program, together with five academic research institutes and seven scientific-production associations. Figure 2 illustrates the expected growth in the global robotics market as estimated by the Tractica consulting firm in 2016. It suggests an exponential growth in sales volume.

1.2 Types of robots and forecasts for their development

In 2017, Internet users could see a dance performed simultaneously by 1,069 Dobi robots (WL Tech) in the city of Guangzhou (South China). According to Mashable.com, the robots' movements were controlled by an integrated system. This show involved a world record-breaking number of simultaneously dancing robots (the previous world record was 1,007 robots). Entertainment robot shows are becoming common practice, boosted by ambitious developers in a rage for new records. Dobi robots are capable of more than mere dancing and singing; they can imitate boxing and kung fu.

Operator-driven robots are nothing more than *machines controlled from a remote desktop*. Such control is easy to perform, even by a child who knows how to play with radio-controlled toys.

Exoskeletons are special devices connected with moving parts of the human body to imitate their movements with enhanced power and improved accuracy. Exoskeletons are finding increasingly wide application in medicine as disability aids for peoples with pathologies of spatial movement. They are also used for remote manipulations of objects in aggressive media and in activities requiring a high accuracy of micromovements, e.g., in microsurgery.

The performance of these devices can be controlled by electric potentials coming to the muscles from the central nervous system, excluding tremors. The potentials are transferred from two nerve endings of antagonistic muscles (flexors and extensors) of hands, legs, and fingers (toes). In this country, V S Gurfinkel' and A E Kobrinskii and co-workers created a biocurrent-driven 'artificial hand' as early as the 1960s–1970s [2, 3]. E B Babskii developed a cardiac pacemaker implanted in the patient's chest cavity and powered by heartbeats via an atrial potential enhancer [4, 5].

Today, robots successfully imitate human movements, recognize individual faces in a crowd, reproduce the voice and facial expressions of a concrete person, and help to design programmed dialogues in the form of determinate question $\leftrightarrow$ determinate answer. Robots can imitate the behavior of a secretary meeting visitors, a waiter taking orders, an exhibition guide, an attendant in a hospital or a senior care home, a lecturer giving pre-programed answers to listeners' questions, a sales person, etc. [6].

Moreover, virtual reality (VR) headsets and mixed surrounding reality–VR headsets operating via optical and acoustic receptors of the wearer in the virtual (artificial) world, as opposed to the real one, are becoming increasingly popular. The former is programed by a computer. The creation of virtual reality is finding application in the design of simulators for training astronauts, pilots, flight dispatchers, surgeons, and other professionals, as well as in project engineering and architecture [7].

Nanorobots and microrobots make up a special class of robotic devices implanted into the human body for the purpose of medical diagnosis or microsurgery, and for the targeted delivery of medications to affected organs (tissues). Microcapsules (nanoplatforms) are currently available that contain biologically active matter or the so-called smartdust exchanging information and merging into clusters for a multitask operation [8]. Another class of nanosystems is represented by nanomachines [9], such as the very dangerous self-reproducing grey goo [10].

The following is just one example from our experience. Attempts to create micro- and nanomachine technology date back more than 40 years [11]. By way of illustration, we realized simple variants of such technologies in the 1980s for the development of a gas-transporting blood substitute based on the ability of 30–70-nm particles of a perfluorocarbons mixture to transfer oxygen from erythrocytes remaining after the loss of blood into tissues. An erythrocyte of arterial blood as large as 100 particles of perfluorocarbon emulsion serves as an oxygen carrier (Sherpa). Perfluorocarbon particles periodically circulate between erythrocytes and tissues in pulsed hydrodynamic plasma flows. They form linear structures resembling 'strings of pearls' that transfer oxygen from erythrocytes to the vascular wall following the Fick law, whence it diffuses into the tissue [12].

As far as *creative android robots* are concerned, the realm of their application in the future is aggressive media or extraordinary and emergency situations, including long-range space flights. CARs are expected to be superior to humans not only in terms of workspeed, accuracy of performing various operations, and power indicators but also in terms of cognitive potential. Although the ability of CARs to self-formulate the goal must be restricted to avoid problems in the robot–human collaboration (see Sections 5.2 and 5.3).

There are many unresolved problems associated with this class of robots. Their development implies the necessity of joint physical, technical, and neurophysiological research. All the processes for which an algorithm can be written are realizable in AI systems. In this context, the following questions arise: "Are there processes that resist mathematical description and algorithmization? Is there a limit to the similarities between CARs and humans?"

In addition, there is another interesting scientific problem concerning our ability to successfully control neurodegenerative diseases. Such control is impossible unless we fully understand how the human brain functions and how humans think. Such understanding would provide a basis for the creation of AI and CARs with human-like intelligence and the formulation of limits on similarities between robots and humans.

In medical practice, schizophrenia is interpreted as the transition of brain work to the chaotic regime. The term *apophenia* was coined in 1958 by the German psychiatrist Klaus Conrade (Greek ἀποφαίνω — expose, reveal) to describe the preoccupation with long-standing psychotic delusional ideas. Such a condition is associated with the early stage of schizophrenia. The human brain in search of solutions necessitated by a difficult situation always acts at the order–chaos borderline, i.e., has to choose between classical and probabilistic logic. A disturbance of cognitive processes in the brain entering the stochastic regime and trying in vain to get out of this state is accompanied by the development of abnormal emotional reactions.

The journal *Physics–Uspekhi* has published many articles with a neurophysiological slant [13–22]. The present review differs from them in that it is focused on the logic of processes proceeding in the human brain and the AI of creative robots.

Forecasting is known to be a thankless occupation, because predictions rarely come true as a consequence of system nonlinearity. Nonetheless, here is one of the variants (Table 1) proposed in 2016 based on the Multidimensional Information Varying Adaptive Reality program complex (MIVAR, Russia), with reference to a variety of expert judgements and propositions made in various years (see the site: www.robotrends.ru).

Table 1. Prospects of development of robotics engineering for the period up to 2040.

Year	Expected result	Reference with the year of prognosis
2018	Robot will pass driving permit test	Oleg Varlamov, MIVAR president, 2016
2019	90% of institutions will have a staff member holding the position of chief data officer (CDO)	Cortner, 2016
2020	30 thousand unscrewed civil aerial vehicles will be in use in the USA	US Federal Aviation Administration, 2012
2022	Robots will be capable of understanding human behavior and responding to it	Expert and analytical report of Rosbusinessconsulting (RBC), 2014
2024	Commercially available motorcars will be capable of reacting to changes in the traffic situation and moving autonomously	IHS Automotive, 2014
2025	Application of industrial robots will reduce labor payment expenditures by 16%	Boston Consulting Group, 2015
2028	First autonomous medical microrobots will be capable of independent directional movements in a patient's body	"A Roadmap for US Robotics: From Internet to Robotics", 2013
2029	AI will be capable of self-learning, understanding jokes, and imitating emotional expressions	Ray Kurzweil, Google's director of engineering, 2014
2030	Commercial androids will have an outer appearance and capabilities identical to those of the human	Expert workshop *Trends and Prospects of Robotics Development in Russia, 2014*
2032	Robots will exceed intellectual potential inherent in humans	Dave Evans, Cisco chief futurist, 2011
2035	In Japan, robots will master 49% of the 600 currently existing professions	Nomura Research Institute, 2015
2040	Law-enforcement robots will exist	Professor Noel Sharkey, University of Sheffield, 2012

1.3 Main differences between robots and humans today

Evidently, existing robots are electromechanical machines, while living organisms represent aqueous-based physicochemical 'machines'. The brain and its functions developed in the course of evolution via *block-hierarchical selection mechanisms* [23, 24] in close connection with the development of the whole organism and environmental changes on our planet. An important role in this process was played by the internal environment of the organism, especially the dialogue between the heart and the brain. This fact if often disregarded when comparing the robot's AI and human intelligence. The heart generates very apparent acoustic and electromagnetic fields, the topography of which on the surface of the body is readily distinguished using modern methods of spatial cardiologic auscultation (Latin *auscultatio* — listening) and electrocardiography.

Acoustic waves are known to be longitudinal elastic oscillations of pressure in gases, liquids, and solids. Each particle in liquids and solids is liable to oscillate around the equilibrium point (standing waves). However, single fluctuations in nonlinear media can produce a solitary wave behaving like a particle. In this case, a special type of wave — the *soliton* — is likely to appear, passing through one another and propagating over large distances [25, 26]. Solitons differ from harmonic waves (see Section 2 for details).

There are still many differences between the robot and the human in both the mode of adaptation to the ambient environment and intellectual creation (Table 2).

The data in Table 2 reflect the current situation. All characteristics of living systems occur separately in nonliving ones [23]. However, the former, unlike the latter, are capable of adaptation to environmental changes. The main prerequisite for this capability is the dynamic stability of the system residing nonequilibrium state [27]. *The processing of information originating from the outside occurs at all hierarchical levels of the organism.*

Table 2. Comparison of the main characteristics of robots and humans.

No.	Main characteristics	Robot	Human	Superiority*
		General characteristics		
1	Principle of organization	Electromechanical machine	Aqueous-based physicochemical 'machine'	—
2	Force exertion	Practically unlimited	Limited	R
3	Fatigability	None	Takes place	R
4	Efficiency	60–90%	15–20%	R
5	Ability to self-formulate the goal	None	Takes place	H
6	Social facilitation**	Possible	Takes place	H
7	Self-association into groups	Achieved in some cases (e.g., drone swarms)	Takes place	H
8	Self-reproduction	Possible	Takes place	H
9	Number of hierarchical levels for processing environmental information	Still limited and fewer than in humans	Limited but greater than in existing robots	H
		Comparative characteristics of artificial robotic and human brains		
10	Characteristic time of reaction formation	Less than 1 μs	~ 0.1 s	R
11	Principle of operation	Discrete	Analog–digital	—
12	Substrate underlying logic elements	Solid passive medium	Liquid, active medium	—
13	Plasticity	Low	High	H
14	Substrate behavior	Stable	Pulsed	—
15	Thermal stabilization	Distributed, mostly air cooling	Local liquid cooling	H
16	Interaction with the system's internal environment	From top to bottom	Hierarchical and cyclic with top–down and bottom–up feedback	H
17	Noise impact	Disturbance in operation	Noise as a creative factor providing state-to-state transitions	H
18	Asymmetric responses to environmental changes	Absent, incapable of creativity	Capable of creativity	H
19	Operation logic	Deterministic (classical) logic	Borderline of classical/probabilistic logics	H

* Notations: R — robot, H — human.
** Facilitation — the mode of control based on a consensus search for assembling system's elements, thus enhancing the stability of the system as a whole. Facilitation process is underlain by a combination of symbiotic and competitive dualism, known in philosophy as "the unity and struggle of opposites".

2. Creativity problem

2.1 Pyramid of hierarchical levels in the organization of living systems

At least nine hierarchical levels are distinguishable in living systems (Fig. 3).

Construction of a mathematical model of a hierarchical system implying unification of all the levels into a whole remains to be accomplished. It seems that results of interactions among only nine levels are not difficult to determine, bearing in mind the small number of all interlevel links given by the relationship

$$Q = m(m-1)\,, \qquad (1)$$

where Q is the number of links, and m is the number of levels; therefore, at $m = 9$, one has $Q = 72$. However, the situation is much more complicated than that, because the number of interacting elements at each level is enormously large due to the fact that they form networks.

Numerous attempts have been made to unite the levels into a single system. For example, the idea of the 'hypercycle' suggested by Manfred Eigen based on the analysis of *macromolecules* dates back to the 1970s [28, 29]. The essence of this notion is illustrated in Fig. 3b: the reproduction of elements in a living system is underlain by template-mediated replication of DNA. This cycle can be labelled by the letter I with a subscript. There can be many such microcycles: from 1 to n.

Autonomous replication within each cycle results in template-to-template rewriting: negative to positive, positive to negative, etc. To ensure mutual regulation of the microcycles, M Eigen proposed that a uniting hypercycle be introduced to facilitate the complication of structural organization in a far-from-equilibrium chemical system with the formation of new microcycles and numerous feedback loops. Each microcycle within the hypercycle accumulates coupling factors (E_i enzymes), simultaneously with replication. These factors can either selectively increase both the rate of replication and its accuracy or decrease the decay rate of already synthesized matrices. This means that any ith microcycle in the hypercycle must always depend on the coupling factor E_{i-1} regardless of what it encodes and rewrites in the replication regime; moreover, it must contain the coupling factor E_{i+1} for a neighbor. M Eigen thought that each ith microcycle serves as a coordinatively controllable genetic unit, and altogether they function sequentially through the hypercycle.

In other words, M Eigen described the phenomenology of the biological phenomenon in question at the macromolecular level, but circumvented all the difficulties of explaining its development and functioning by introducing a specific 'guiding principle' in the form of coupling factors E_i that 'can do all things', while becoming increasingly more complicated. However, he failed to clarify the kinetics and driving forces of such self-complication. In a word, the process of hypercycle formation remains unexplained, even if the idea of cycles can be regarded as productive.

In fact, Eigen's hypercycle is the projection of higher pyramid levels (Fig. 3a) onto the macromolecular level. Such a reduction results in the loss of the specific kinetics of the processes taking place at the remaining hierarchical levels.

Another idea was forwarded by M I Rabinovich et al. [14, 30]. In what follows, it will be clear that this idea is close to that considered in the present review; specifically, it is an attempt to create a mathematical model of consciousness based on the vibration theory in terms of nonlinear equations. However, the approach proposed by Rabinovich and coworkers disregards the role of '*good luck*' in the ensemble of a large number of degrees of freedom involved in explaining the spontaneous emergence of order from chaos. In psychology, 'good luck' is often treated as *enlightenment* (see the parable of the bright child in Section 5.2). In this regard, *chance events constitute one of the important mechanisms underlying the formation of inter-neuronal connectivity*, while nonlinearity

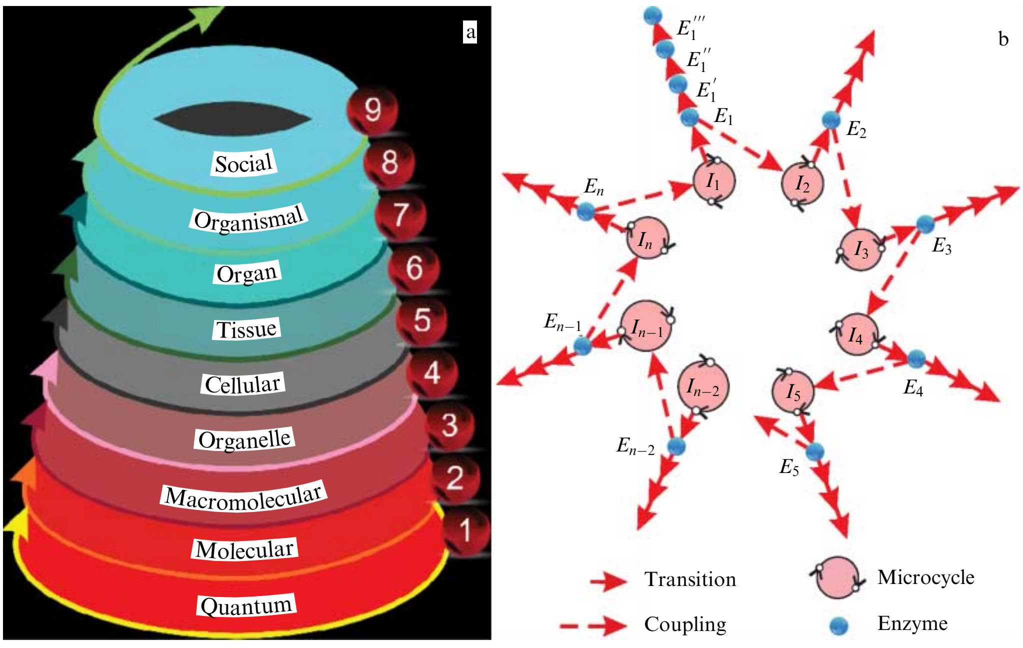

Figure 3. Hierarchical organization of living systems: (a) organizational levels, and (b) M Eigen's hypercycle.

acts at a later stage to prevent unlimited growth. Nerve impulses generated in interactions with the external environment are apparent as singular peaks above the noise level [31]. Intervals between the impulses can be rather long. The general name of such a picture in physics is 'intermittence' [32]. The problem of the origin of intermittence is akin to that of the recovery from determinate chaos to ordered motion in a dynamical system [33].

There are only two requirements to any proposed model, viz. internal consistency and predictability of new regimes confirmed in experiment. Hence, there are a great variety of models meeting these criteria for the description of complex systems (see Section 3).

R Penrose wrote in his book [34, p. 14]: "*Inherent in the formation of our consciousness are elements that cannot be derived from any set of computational instructions.... It is impossible to find place for 'nonalgorithmic actions' in the framework of universally accepted physical theories. Therefore, we must look for a gap in the scientific picture into which such actions might be inserted. I claim that this 'blank spot' lies somewhere at the borderline between 'submicroscopic' world of quantum physics and macroworld obeying classical mechanics laws*".

Giving credit to Penrose for these thoughts, it is opportune to mention certain causes accounting for significant differences in information processing and the form of the energy maintaining it between human and robotic brains. There is the only form of energy needed to enable robotic operations: electric energy, whereas the energy of human brain (and of the whole organism) is included in the energy transformation cycle on our planet: the Sun → the Earth with its atmosphere ↔ flora ↔ fauna. Humans are a component of the biospheric food web and cannot exist outside it.

We substantially extended the dimensions of the observable world with the advent of microscopes, telescopes, accelerators, fast means of transportation, and external memory systems (languages, writing, arts, Internet, etc.). But the number of natural hierarchical levels in the human body + brain system where information received from the external world is processed, remained practically unaltered. Notice that living systems also have a quantum hierarchical level built up by photosensitive proteins (chlorophyll and rhodopsins) and ionic, molecular, and macromolecular levels. Higher organizational levels of life are represented by organelles, cells, tissues (networks), organs, the whole organism and, finally, the social level at which individuals collaborate, making use of acoustic, olfactory, and body languages perceived by hearing, olfaction, and sight.

The main cause of diversity of hierarchical levels in living systems appears to be the great number of different guises of the surrounding world, from an absolutely determinate one governed by Newton's laws to the absolutely chaotic probabilistic Einstein–Smoluchowski (big billiard) world [35, 36] or a combination of both (chaos or intermittence, determinateness or stochasticity) [32]. Orientation under conditions of continuous environmental changes needs a complementary multilevel organization of information processing capable of rapid adaptation to the continuously modified environment that enhances the stability of living systems. Each action performed by an organism has its price H. The price of the action is the function

$$H = f(E, \tau)\,, \tag{2}$$

where E is the energy consumed, and τ is the time of action.

Under stressful conditions, the energy consumed increases stepwise, while time of action τ_i at each stage decreases; therefore, the frequency Ω_i of performing all actions taken together increases during the upward transition from one level to another. Given that there are N hierarchical levels, each contributing to energy consumption, the overall energy expenditure increases and reaches a maximum. Hence, the expression

$$E_{\max} = \sum_{i=1}^{N} H_i \Omega_i\,, \tag{3}$$

where Ω is the frequency of action. This expression describes a situation under high stress. The total energy consumption in expression (3) is determined by the ith level with the highest item, i.e., by low levels operating at high frequencies (quantum, molecular, and macromolecular levels) (see Fig. 3a).

Under normal low-stress conditions, e.g., learning, the total action price reduces to a minimum, and all N levels operate at half-maximum efficiency as described by the expression

$$H_{\min} = \sum_{i=1}^{N} \frac{E_i}{\Omega_i}\,. \tag{4}$$

In this case, the greatest contribution to the price of action H comes from the levels operating at a low frequency, such as social (public opinion and encouragement), organism (comfort of an individual), and organ (comfort in the internal environment achieved in interactions among organs).

This means that the stability of a system depends on the feedback between upward and downward transitions. Then, *acceleration of environmental changes under the effect of various factors, including human-made ones, must not exceed the acceleration of the adaptation of living systems to these changes*. Otherwise, a system will have no time to adapt itself to them, undergo degradation, and die.

2.2 Dialogues at different levels of hierarchical organization

Human life can be likened to a transient continuous branching process of given duration that requires the correct choice of the direction of motion in time and space with an estimation of the accompanying risks. They increase in number and seriousness under conditions of uncertainty, affecting the accuracy of determining the probability of both positive and negative outcomes. For example, incorrect assessment of such probability in a predator–prey system is fraught with danger and can cause the death of one of its members. The perception of risks by a living organism is based on the feeling of fear genetically inherent in it in the form of an instinct giving rise to either an unconscious or a partly or totally (with the participation of the brain) conscious choice between the decision to fight or flee, agree or contradict.

Inherent in a real organism, including the brain, is a dualism simultaneously clarifying itself in the form of symbiosis (cooperation) and competition (antagonism) between its different parts (cells and organs) and the environment, based on positive and negative feedback. The notion of 'reverse afferentation' (from the genitive case of the Latin *afferentis* — conveying) appears to have been introduced by P K Anokhin in the early 1950s [37]. The synonym

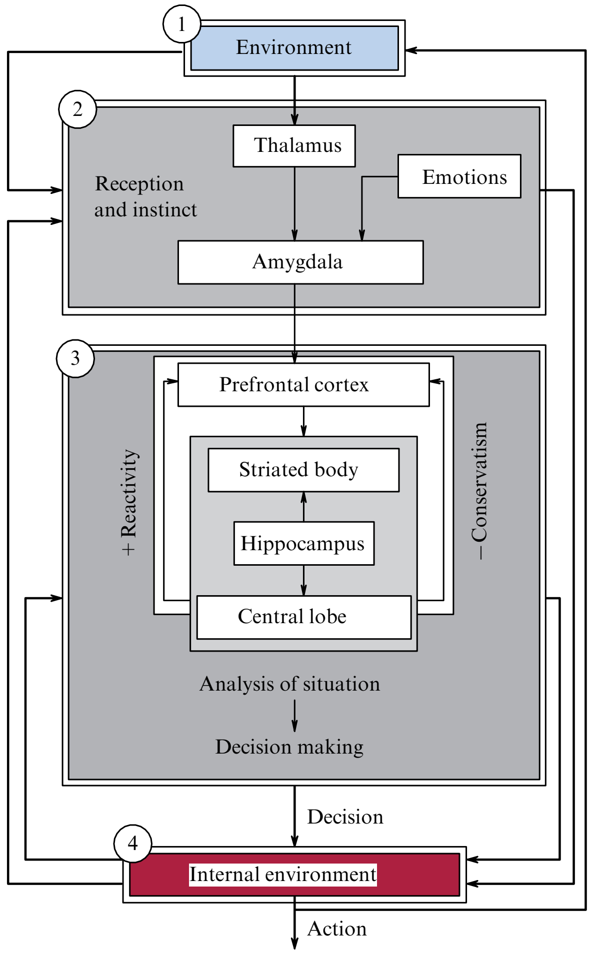

Figure 4. Simplified flow chart of pathways at the level of brain structures with direct couplings and feedback involved in human decision-making. Any decision taken is influenced by emotions and the result of competition between brain structures (striated body and central region). In individuals prone to taking risk and making quick decisions, the striated body is more active than the central region. In slow decision-making people, the central region is more active than the striated body. The compromise between them taking into account the influence of environmental factors and past experience (hippocampus-controlled memory) creates a stable state and generates a signal for action.

of this term in cybernetics is retained under the technical name 'feedback'. To recall, this notion in dialectics has been viewed as the law of the 'unity and struggle of opposites' since Hegel's time. A result of such a struggle is the 'removal of contradictions' [38]. All the systems of a living organism, including different brain regions, operate via feedback mechanisms that form cycles and thereby remove inconsistencies to maintain competition between them within the set bounds. An example is presented in Fig. 4.

However, such competition, either dispute or dialogue (the name is a matter of taste), occurs at different organizational levels of a living system in different languages: biochemical, electrical, or acoustic. Antagonistic interactions between hormones are especially well explored at the biochemical molecular level (e.g., between the fear hormones adrenalin and noradrenalin, sometimes called rabbit's hormone and lion's hormone, respectively, in accordance with their influence on the organism). The purpose of these dialogues is to reach an agreement.

To begin with, reaching the agreement at each level guarantees the stability of the organism as a whole. Second, it leads to compression of information, which facilitates, if successful, remembering it for future use. Moreover, it saves storage capacity for constructing a virtual model of the outer environment in the brain. As a result, information in living organisms is transferred between levels and undergoes transformation, for instance, owing to links formed between upper and lower levels it passes from the brain to the spinal cord and assumes the function of motor skills to free the upper levels from routine work [39]. This additionally allows saving energy and reducing entropy within a relatively short span of time. Information transfer from one level to another can influence the behavior of individuals in social systems. For example, a human adapts to the behavior of a crowd [40], and an animal to the behavior of a herd or a flock [41–43].

2.3 Dialogue between the heart and the brain

This situation at the organ level is exemplified by *a dialogue* between the heart and the brain. The beating heart 'talks' with the brain in an acoustic language by emitting infrasounds with a frequency from 0.75 to 2.5 Hz (on average, ~ 1 Hz) into the intracerebral fluid and cerebrospinal fluid that contains hormones and energy sources (sugars and oxidizing agents). Movements of the liquid result in intermixing. The sound propagation distance depends on absorption by irregularities of the medium, refraction (bending of sound rays in an inhomogeneous medium), and scattering. In addition, refraction is manifested the stronger, the greater the sound speed gradient. The sound propagation distance increases as the acoustic vibration frequency drops, and soliton-like movements are likely to develop [25, 26]. Sound waves reflected from irregularities can merge, thereby extending the rear-front of an acoustic pulse and prolonging its action.

Hormone distribution in different brain regions is responsible for the 'emotional coloring' of perception of the environment. The brain answers the heart in the language of electric signals propagating along axons and dendrites that, in turn, alter local density of the fluid by almost 70%. Changes in density are due to a surge in potassium ions in response to neuron excitation. The mass of potassium ions being higher than that of sodium ions, local density of the medium around the axons increases. Simultaneously, the reflection of sound waves becomes more pronounced, and their velocity in the denser medium increases.

Furthermore, the brain controls heart rate by sending signals to the sinoatrial node via a feedback loop (1 in Fig. 5a). The heart, in turn, influences brain work via pathway 2 (Fig. 5b).

The speed of sound in *water* is around 1300–1500 m s^{-1}. An excitation pulse propagates in axons at a much lower speed (around 25 m s^{-1}).

Even at rest, roughly 15% of the total blood volume passes through the brain, which consumes up to 20–25% of the oxygen brought in with inhaled air [44]. Blood is supplied to the brain through the internal carotid arteries and two vertebral arteries. The posterior and anterior vertebral arteries form a circular blood vessel (Fig. 5c). Venous outflow occurs mainly through two jugular veins.

The role of the heart in its dialogue with the brain is seemingly insignificant, being reduced to delivering nutrients and disposing the useless or harmful metabolites. However, it

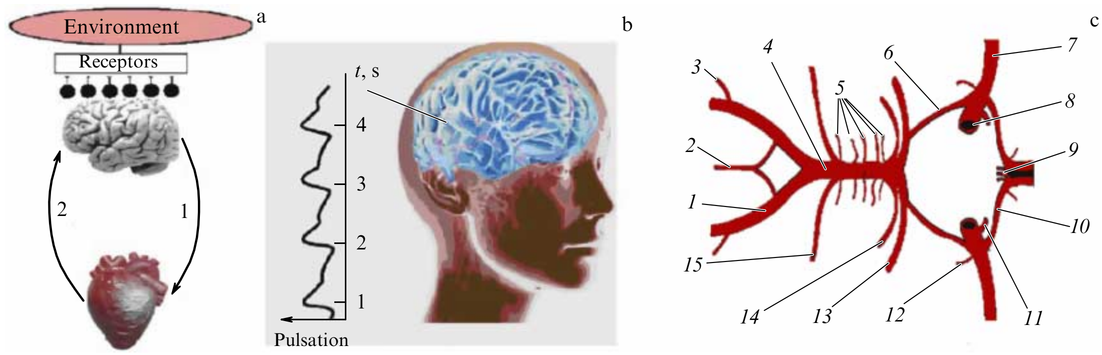

Figure 5. Heart–brain interplay in the volume of cerebrospinal fluid (liquor). (a) Diagram of interaction. (b) CSF volume fluctuations in time caused by heart contractions. (c) Major brain-feeding arteries: *1* — vertebral artery, *2* — anterior spinal artery, *3* — posterior inferior cerebellar artery, *4* — basal artery, *5* — arteries of pons, *6* — posterior communicating artery, *7* — middle cerebral artery, *8* — internal carotid artery, *9* — anterior communicating artery, *10* — anterior cerebral artery, *11* — principal artery, *12* — anterior choroidal artery, *13* — posterior cerebral artery, *14* — superior cerebellar artery, and *15* — inferior cerebellar artery. Veins are not shown.

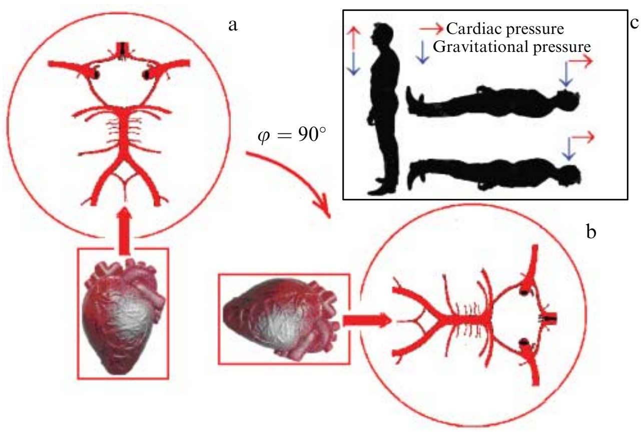

Figure 6. Postural change alters the relative position of the heart and cerebral blood system (a, b) with respect to the gravitational pressure vector (c). The brain functions only at a relatively constant difference between intracranial and environmental pressures and temperatures.

is not so. The heart, in cooperation with the brain's glial network, redistributes blood flows both at the macrolevel and inside each portion of the neural network. Such redistribution promotes adaptation of the organism to environmental changes and influences in stressful situations the rate of decision-making through the adequate supply of proper hormones by the brain.

The outer environment exerts influence on the human body and internal organs via atmospheric pressure and gravitational pressure. The brain changes the heart rate to make up for these effects. The atmospheric pressure field varies within 6%, regardless of human body posture. In contrast, gravitational pressure with respect to the blood pressure vector is directed largely along the body axis and depends on the posture (Fig. 6).

The processes in the brain fluid can be described in terms of different variants of the laws of conservation [44–46]. The cerebrospinal fluid (frequently termed *liquor*) contains a great variety of compounds and ions. Its average density is close to the density of water: 1.005–1.007 g ml^{-1}, the difference being only 0.5–0.7%. It can thus be regarded as an ideal incompressible liquid, and we can use Poiseuille's law to determine its viscosity μ. In the upright position of the human body, one finds

$$\mu_1 = \frac{\pi r_{\mathrm{c}}^4 (p_1 - p_2)}{8 Q_1 l}\,, \tag{5}$$

where r_{c} is the capillary diameter, p_1 is the cardiac pressure, p_2 is the gravitational pressure, Q_1 is the fluid flow rate in the vertical position, and l is the length of capillaries.

Vector summation of pressures in the horizontal human position gives the coefficient of viscosity

$$\mu_2 = \frac{\pi r_{\mathrm{c}}^4 (p_1^2 + p_2^2)^{1/2}}{8 Q_2 l}\,. \tag{6}$$

A change in the posture leads to the relationship

$$\frac{\mu_1}{\mu_2} = \frac{Q_2 (p_1 - p_2)}{Q_1 \sqrt{p_1^2 + p_2^2}}\,. \tag{7}$$

If viscosity remains unaltered, for linear laws at a local portion, one has

$$\frac{\mu_1}{\mu_2} = 1\,. \tag{8}$$

It follows from formulas (5) and (6) that

$$Q_1 (p_1^2 + p_2^2)^{1/2} = Q_2 (p_1 - p_2)\,, \tag{9}$$

which means that the flow rate of a fluid with constant viscosity grows in inverse proportion to pressure upon a change in human posture. There are two variants of regulation in the case of a postural change. One is the influence of the brain on the heart mediated through the feedback loop resulting in a decrease (increase) in heart rate p. The other is blood flow mass control by vascular constriction or dilation effected by the system composed of the *heart → blood vessels → glial cells → and neurons*.

The fluid flow velocity and therefore the brain pulsation rate differ in different parts of the system. Flow line density is consistent with the velocity in a given part. The *stationary* or *steady-state flow* corresponds to the *continuity equation* describing the passage of a given fluid volume through any distinguished cross section for equal time intervals. Let S_1 and S_2 be two cross section areas, and $\mathbf{u}_1$ and $\mathbf{u}_2$ the respective velocity vectors of fluid particles in the uniformly moving flow. Then, the continuity equation has the form $\mathbf{u}_1 S_1 = \mathbf{u}_2 S_2$. It holds for all flows, i.e., $\mathbf{u}S = \text{const}$. However, one more dynamic variant of fluid behavior is conceivable when 'banks' pulsate with frequency ω_1, and fluid volume with ω_2. Then, the new invariant relation will take on the form

$$\mathbf{u}_1\omega_1 = \mathbf{u}_2\omega_2 \quad \text{or} \quad \mathbf{a}_1 = \mathbf{a}_2 \,, \tag{10}$$

where $\mathbf{a}_1$ and $\mathbf{a}_2$ are acceleration vectors of the two flows. The total liquor volume in a healthy adult human ranges from 140 to 270 ml, accounting for roughly 20% of the brain's weight. Let us consider in more detail the fluid motility dynamics at density ρ_1 in a semisphere.

Suppose that the semisphere is partly (e.g., 20%) filled with the fluid. Then, the fluid height $h = f(R)$ is restored at any position of the semisphere in space for the characteristic time of the transient process due to fluid translocation. Brain structures either slow down or speed up the translocation depending on the relative direction of acceleration vectors. The mass of the fluid thus transferred is proportional to the volume V_s of the semisphere it fills:

$$V_s = \frac{1}{6}\pi h(h^2 + 3r_1^2) = \pi h^2\left(R - \frac{1}{3}h\right), \tag{11}$$

and to the fluid density ρ_1, where h is the height of the fluid in the semisphere, r_1 is the radius of the segment surface, and R is the radius of the sphere. Taking into account expression (11) yields the fluid mass

$$m = \rho_1 V_s = \pi\rho_1 h^2\left(R - \frac{1}{3}h\right). \tag{12}$$

Angular momentum L is expressed as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}\,. \tag{13}$$

Here, $\mathbf{r}$ is the radius vector originating in the center of gravity of the brain mass, and $\mathbf{p}$ is the vector of the impulse of force. Differentiation of expression (13) yields

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} \times \mathbf{p} + \mathbf{r} \times \frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{u} \times \mathbf{p} + \mathbf{r} \times \mathbf{F}_{\text{res}}\,, \tag{14}$$

where $\mathbf{u}$ is the velocity vector, and $\mathbf{F}_{\text{res}}$ is the resulting force vector. In a vertical position, the product of $\mathbf{u} \times \mathbf{p}$ is zero, because vectors $\mathbf{u}$ and $\mathbf{p}$ are parallel to each other. Similarly, the term $\mathbf{r} \times \mathbf{F}_{\text{res}}$ vanishes as well, since central forces created by the heart are parallel to vector $\mathbf{r}$. Hence, we put

$$\frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t} = 0\,, \quad \text{or} \quad \mathbf{L} = \text{const}\,. \tag{15}$$

A change in the posture does not alter the first gravitational term ($\mathbf{u} \times \mathbf{p}$) on the right-hand side of Eqn (14) that equals zero. The second term is the vector sum of gravitational and cardiac forces. The resultant moment of force $\mathbf{T}_{\text{res}}$ is expressed as follows:

$$\mathbf{T}_{\text{res}} = \frac{\mathrm{d}\mathbf{L}}{\mathrm{d}t}\,. \tag{16}$$

This means that the resultant moment of force is equal to the rate of change of the angular momentum. If all brain structures are assumed to be closed, $\mathbf{L}_{\text{res}} = \text{const}$ as a direct consequence of Newton's laws. It allows the equation of blood flow acceleration for the two extreme postures (vertical and horizontal) to be written out in three variants. Let us denote blood flow acceleration in the vertical and horizontal positions by $\mathbf{a}_\uparrow$ and $\mathbf{a}_\rightarrow$ respectively. Then, the corresponding resultant forces in the two respective positions are $(F_{\text{heart}} - F_{\text{grav}})_\uparrow$ and $(F^2_{\text{heart}} + F^2_{\text{grav}})^{1/2}_\rightarrow$. There are three cases in the dependence on the acceleration ratio:

$$\begin{aligned}
&(1)\ \ \mathbf{a}_\rightarrow > \mathbf{a}_\uparrow \quad \text{for} \quad \frac{F_{\text{heart}} - F_{\text{grav}}}{\sqrt{F^2_{\text{heart}} + F^2_{\text{grav}}}} > 1\,,\\
&(2)\ \ \mathbf{a}_\rightarrow = \mathbf{a}_\uparrow \quad \text{at} \quad \frac{F_{\text{heart}} - F_{\text{grav}}}{\sqrt{F^2_{\text{heart}} + F^2_{\text{grav}}}} = 1\,,\\
&(3)\ \ \mathbf{a}_\rightarrow < \mathbf{a}_\uparrow \quad \text{for} \quad \frac{F_{\text{heart}} - F_{\text{grav}}}{\sqrt{F^2_{\text{heart}} + F^2_{\text{grav}}}} < 1\,.
\end{aligned} \tag{17}$$

Five conclusions follow from the consideration of these three cases:

(1) Case 1, when $F_{\text{heart}} > F_{\text{grav}}$ is a normal state.

(2) Case 2, when $F_{\text{heart}} \gg F_{\text{grav}}$ corresponds to a significant relative decrease in gravity (weightlessness). This condition is dangerous, since it can result in a stroke. To return to the normal state, the brain has to reduce the heart rate or increase the amount of fluid, while the $(F_{\text{heart}} - F_{\text{grav}})/(F^2_{\text{heart}} + F^2_{\text{grav}})^{1/2}$ ratio tends toward unity, because

$$\lim\left(1 - \frac{F_{\text{grav}}}{F_{\text{heart}}}\right) \to 1 \quad \text{as} \quad F_{\text{heart}} \to \infty\,.$$

(3) Case 3, when $F_{\text{heart}} < F_{\text{grav}}$ corresponds to a pathological condition, e.g., cardiac failure, which makes the affected subject unable to maintain an upright position due to cerebral venous outflow leading to the loss of consciousness.

(4) Finally, the special case is hypergravity, $F_{\text{heart}} \ll F_{\text{grav}}$, when

$$\lim\left(\frac{F_{\text{heart}}}{F_{\text{grav}}} - 1\right) \to -1\,.$$

This dangerous condition can also cause the loss of consciousness. It accounts for the fact that astronauts being prepared for a space mission consider training in a rotating chair or centrifuge to be a most trying experience.

The main conclusion is formulated as follows: *body fluids are always in motion and can travel with acceleration under the influence of regulatory systems, including the brain. This mechanism makes up the basis for one of the variants of environmental adaptation. In this case, the adaptation is a way to maintain functioning of the organism when the position of the body in a three-dimensional space changes in the presence of gravity*. Robots do not have such an adaptive mechanism, nor do they need it. It can be created based on other technical principles, whenever needed in certain exotic cases.

2.4 Simple model for reaching consensus

Our brain is not only a 'referee' clearing up controversial situations arising from the competitive relationships between human internal organs using feedback but also an interface that correlates genetically determined instincts underlying bodily needs and their satisfaction under varying environmental conditions (Fig. 7a).

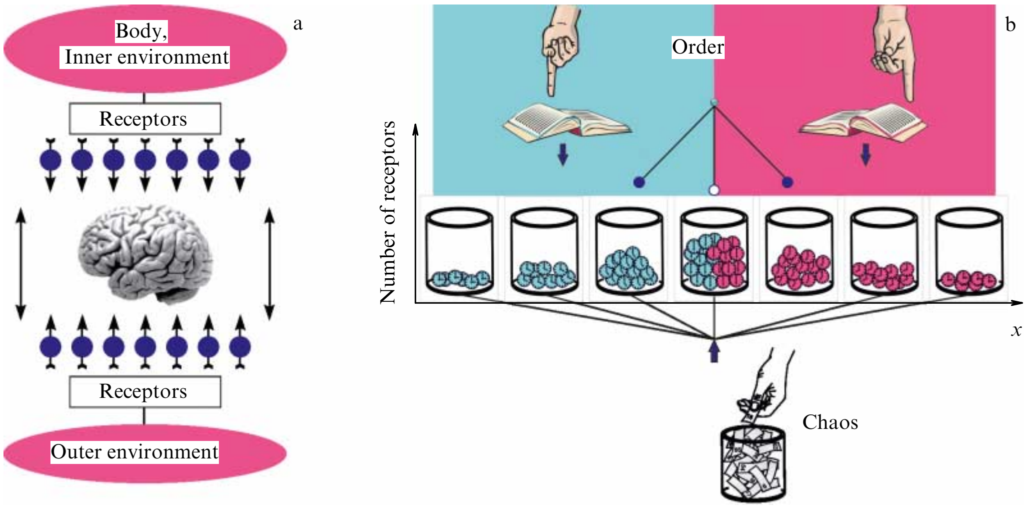

Figure 7. Search for homeostatic equilibrium between the inner environment and ambient changes: (a) the brain as an interface correlating the inner and outer environments; (b) the simplest computer model of chaos–order transition based on two populations of competing receptors.

Our receptors frequently perceive the outer environment as a random sequence of events (e.g., determinate chaos) [47]. The brain tries to cope with such situations by taking advantage of competition among different receptor species. Let us consider the correlated actions of receptors in the framework of the simplest computer-assisted *conceptual machine model* that we used to study the dynamics of the random behavior of a complex system composed of N sensors interacting in accordance with periodically changing specified rules under external control conditions [48, 49]. The process of interest is schematically presented in Fig. 7b. The system operates in the form of periodic chaos $\leftrightarrow$ order transitions.

Let us describe a dialogue model with two possible results in the limit: a double-pole variant (futile dialogue) and a single-pole dialogue resulting in reaching the consensus. Let us further assume that there are K vessels, each with n_i sensors interacting among themselves (Fig. 7b). The total number of sensors is

$$N = \sum_{i=1}^{K} n_i \,. \tag{18}$$

Suppose for simplicity that $N = \mathrm{const}$, i.e., consider the system within a small time interval, when replication can be disregarded. Each sensor has its own number and properties. The initial properties of the sensors are plotted on the x-axis in Fig. 7b. Their numbers are intermixed in a vessel marked Chaos. Two numbers are randomly taken from the Chaos vessel on a periodic basis with frequency Ω, which symbolizes the encounter and interaction of two sensors previously located on the x-axis in their host vessels, either different or the same. Interaction rules for each pair of sensors can change to the opposite depending on the control system, i.e., the position of an external pendulum oscillating with frequency ω and setting the interaction rhythm in the capacity of a biological clock:

$$\Omega > \omega \,, \tag{19}$$

where Ω is the frequency of encounters between sensors. When the pendulum is on the right, the opinion (property) of the sensor from the right vessel is assigned to both encountering sensors. Conversely, if the pendulum swings left, the opinion of the sensor from the left vessel is assigned to the two sensors. The properties of both sensors initially located in the same vessel remain unaltered when they encounter each other. After each encounter, the sensors are returned into the Chaos vessel, where they merge. As a result, the sensors involved in pairwise interactions as described above tend to change the summarized opinion distribution curve in time and thereby give the answer to the question: "How does the virtual model synthesized by different brain regions change in dynamics when altering sensor opinion distribution upon a rise in the number of encounters?"

In fact, the model being considered demonstrates the dialogue of two disputing groups having different opinions about external and internal situations. Figure 8 represents time-related changes in views as a transition from initial to final distribution during an interval determined by the number of interactions.

It is difficult to analytically describe this probabilistic process, because the number of encounters between sensors having identical opinions is uncertain. Evidently, shaping opinion is a convergent process complicated by uncertainty in transient process duration, i.e., the time interval needed to reach a stable or quasistable (vibrational) state.

The process may stop in two cases: first, the time of the encounter is exhausted before a full consensus is reached; second, the consensus is reached within the specified time, i.e., all differences of opinion are resolved. In the latter case, all the sensors are gathered in one vessel.

The pendulum oscillation period (i.e., a change in circumstances in the external or internal environment analyzed by the brain) is of importance. The oscillative control of a slow-swinging pendulum does not have an appreciable effect on the process. The system has enough time to enter the stable state within one (two or three at the most) oscillation periods, retaining the opinion of the ensemble whose advantages corresponded to the pendulum position in the respective cycle.

Of special interest are cases in which roughly equal time (much shorter than needed to reach a full consensus) is provided to realize each rule. Under such conditions, the system enters a lasting oscillatory regime without an obvious final result.

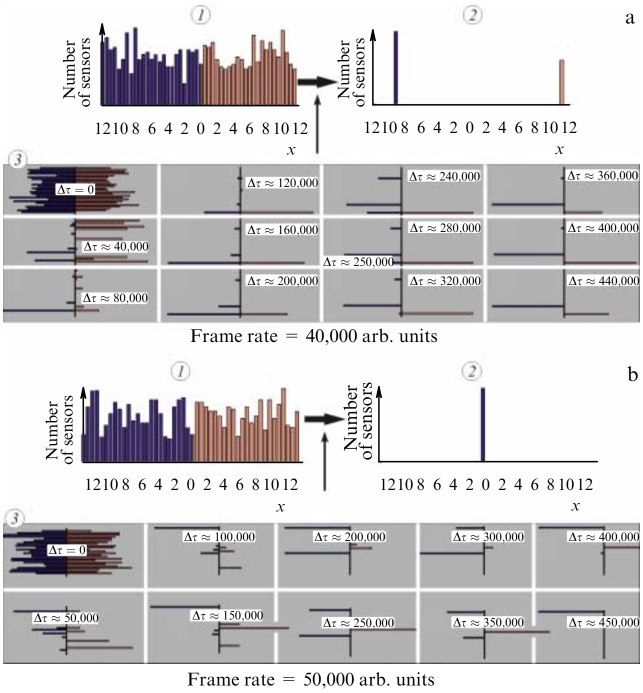

Figure 8. Frame flow chart illustrating dynamic changes in the opinion of two initially different groups of competing sensors (500 sensors per group) in pairwise interactions. Initial properties of the sensors are plotted on the x-axes. Initial (1) and final (2) normalized distributions of opinions are shown at the top of panels a and b. Filmograms (3) demonstrating kinetics of transient process with specified frame rate are presented at the bottom.

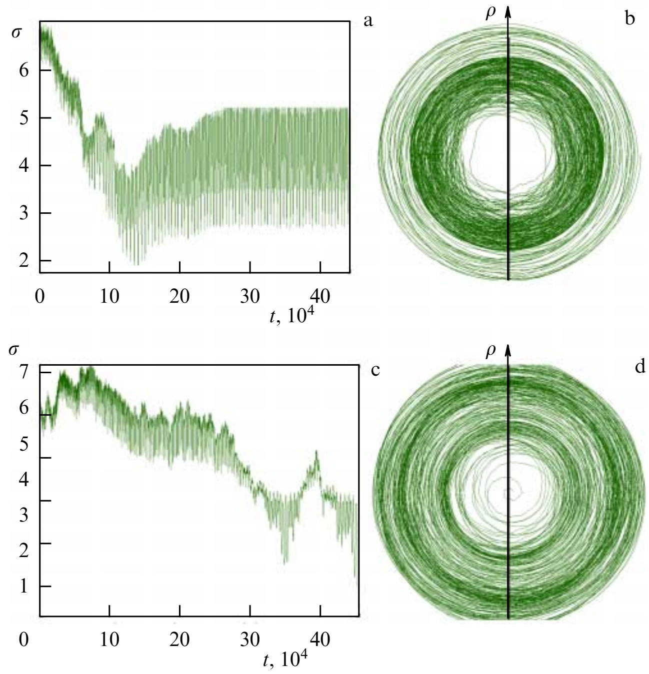

Figure 9. Time-dependent variations in root-mean-square deviation σ over an entire ensemble of sensors. (a) Dynamics of opinion convergence in time for two initially different competing groups of sensors corresponding to the filmogram in Fig. 8a. (b) The same process in polar coordinates, where $\rho = (\sigma^2 + t^2)^{1/2}$ and angle of rotation $\varphi = \arctan(\sigma/t) = \arcsin[\sigma/(\sigma^2 + t^2)^{1/2}]$. (c) Dynamics of opinion convergence in time corresponding to the filmogram in Fig. 8b. (d) The same process in polar coordinates. Because random interactions between sensors under conditions of the model occur at equal intervals, the number of interactions is equal to the past time.

The analysis of the model revealed a sharp rise in the duration of competition between two newly formed identical clusters, giving rise to opposite opinions. The distribution amplitude changes with the pendulum oscillation frequency. The pendulum swinging to the left increases the amplitude of opinions of the left ensemble with its corresponding decrease in the right one. The situation reverses for rightward deflection of the pendulum.

The dynamics of the transient process can also be illustrated by variations in root-mean-square deviation $\sigma(t)$ of the opinion distribution over the ensemble of all sensors in time (Fig. 9).

It was revealed that the model has a number of interesting peculiar features. There is a quite apparent dependence of opinion synchronization time on the symmetry of the initial state and the scatter of gradations in the starting distribution (i.e., genetics). In the absence of asymmetry and local gradation scatter of the distribution, the probability of forming two identical clusters and a process with a long-lasting transient regime increases.

The 'survival' of gradations also depends on their local surrounding. It accounts for the frequent appearance of gradation in distribution initially differing from its maximum value in the starting distribution, because sensors can move into gradation with the initially smaller amplitude, when the rules change. Metaphorically speaking, *any circumstance can moderate the ambitions of the initially leading group of sensors and disprove them*. Random movements of the sensors from one gradation to another follows from the casual description of the encounter process (the Chaos block in Fig. 7b).

Time-dependent changes in radial density in the ring disposition on the scatter plots in polar coordinates suggest a nonuniform convergence to final results of synchronization in the transient regime (Fig. 9b, d).

The time a system needs to reach synchronization under the influence of an external periodic impact (pendular control) increases by several orders of magnitude (e.g., it changes from 10^5 to 10^9 in arbitrary units, when the total number of sensors $N = 10{,}000$). A periodic change in the rules causes a transition of the system into the forced oscillation regime.

Two variants of self-organization produced different results. One is the two-pole oscillatory situation (Fig. 8a), and the other the single-pole statically stable situation (Fig. 8b).

This model describes competition characteristic of paired interactions. At each hierarchical level of a living organism, such interactions are distributed in space and time over the entire body volume, including the brain. Consensus is reached in certain parts of the body and the brain, whereas desynchronization increases in others. Transitions spread like waves from one part to another and return. Such a migration of excitations can be observed in the brain using magnetic resonance imaging (MRI).

2.5 Examples of paradoxes of sensor interactions

One of the problems solved by the human brain is the comparison of information coming through different channels from sensory organs. The brain ignores some data but strengthens the utilization of others. Paradoxes embarrass it,

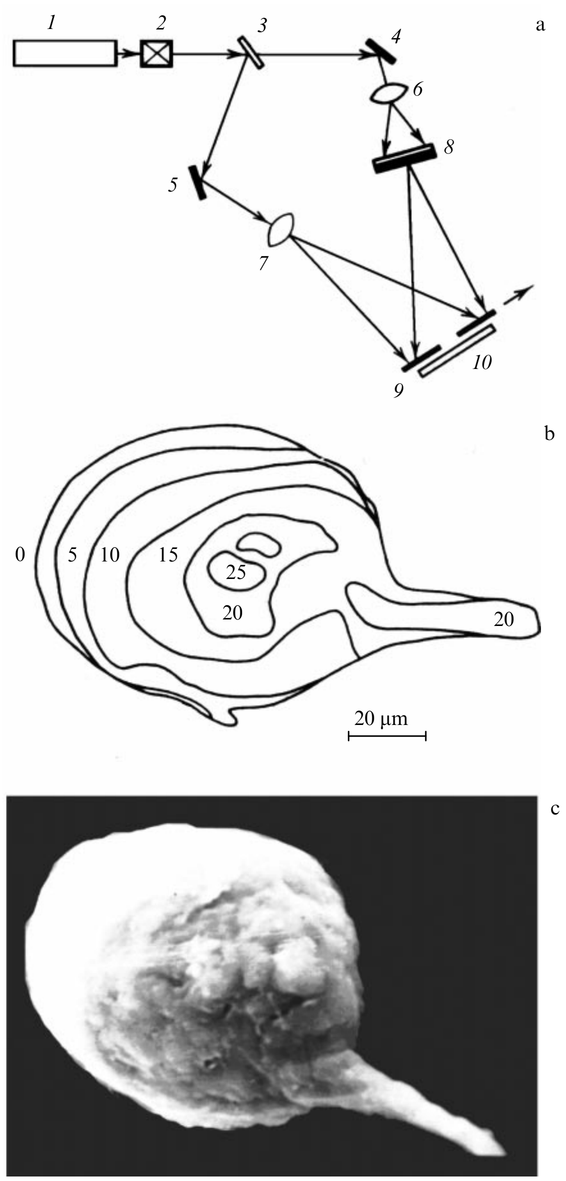

Figure 10. Three-dimensional holographic image of a neuron of the great pond snail (*Limnaea stagnalis*) emerging in the smoggy air as a result of the scattering of light passing through a holographic plate: (a) Schematic of the device for producing composite holographic images based on the photographs taken from different perspectives: *1* — laser, *2* — shutter, *3* — beam splitter, *4, 5* — mirrors, *6, 7* — lenses, *8* — angle shots, *9* — movable slot, and *10* — hologram. (b) Equally high lines in the object (μm) measured with the aid of a light spot inserted into the reconstructed image. (c) Photograph of the hologram of the reconstructed image.

as exemplified by the perception of holograms. In the mid-1970s, we learnt to create 3D holographic images from a set of electron micrographs (Fig. 10) [50–52]. At that time, many people were unaware of the properties of holograms.

A human (or monkey) seeing a holographic image for the first time finds that a hand easily passes through it. The incorporeity of a three-dimensional object may arouse admiration, horror, curiosity, or aggression in the observer [50].

One type of reaction to a hologram was reported in New York newspapers in November 1972 when traffic on Fifth Avenue was jammed for an hour by a crowd gathered near a shop window to see a female hand stretched out of the half-darkness and displaying a diamond necklace. The mysterious holographic image was formed by a powerful laser based on a hologram generated by McDonnel Douglas Electronics Company. The image looked so much like a real object that someone broke the window glass with an umbrella, stating that the whole thing was the devil's work. Our studies showed that the discrepancy between perception through visual and tactile channels disappears as soon as the observer understands (due to correction of the virtual model of environment in the brain) that the hologram is a specific kind of images [51].

Here is another example. Images of all visible objects are known to be projected and rotated 180° by the eye's lens onto the retina. Therefore, a newborn baby needs a rather long time to gain the necessary experience through trial-and-error procedure as he/she tries to take hold of toys suspended over the cradle. The agreement between visual receptors perceiving information and tactile sensations is reached after the brain forms a program for turning the image of an object through 180°.

There are scores of examples of temporal discrepancies arising from the perception of environmental objects with subsequent restoration of agreement between different types of receptors. Suffice it to mention that even three-hour jet lag leads in many people to sleep disturbance and performance impairment. Astronauts have to adapt themselves to the new forms of motion in a gravity-free space during the first days after arrival at an orbital station. Experiments on long-term adaptation to wearing inverting spectacles have shown that more than 10 days are needed for normal spatial orientation to be restored. Traveling along a meridian from one region to another within the same time zone also requires adaptation to new microflora, microfauna, water quality, and other factors. The period of adaptation is significantly different in various human populations; in individuals, it depends on the age. Sometimes, adaptation is accompanied by diseases, such as desynchronosis, allergy, diarrhea, and seasickness. All these adaptive events are mediated through the phase of temporary chaotic activity of the brain with subsequent restoration of harmony among the sensors.

2.6 Virtual model synthesized by our brain

The relationship between the brain's virtual model and human behavior, as well as the important role of this model, is confirmed by numerous facts. For example, *a priori* beliefs may facilitate the formation of an image of fear in a child's mind. If a child remaining alone in a dark room anticipates danger, it evokes the feeling of fear interpreted by the brain as if coming from the eyes, ears, and other sensors, even though none of them send respective signals. The brain creates an image of the fear having size, strength, shape, and other attributes. Most kids are afraid of the dark for the lack of life experience (hence, the Russian proverbial expression "Fear has big eyes").

On the one hand, the boundedness of an outer environmental situation model is of little consequence, because the model in question — like any other model — is a simplified version of reality. It fails to take account of many aspects of the outside world but adequately reflects the gist — that is, an urgency to avoid catastrophic situations.

On the other hand, like many living creatures, humans develop in the course of ontogenesis the evolutionarily determined quality of *curiosity* opposite to fear. In physics, 'the struggle of opposites' has been understood since Plato' time in the sense of dispute surrounding the *antipodes* (ancient Greek ἀντίπους[2] — prefix meaning opposite + ὄνομα — name). In linguistics, antipodes are termed antonyms, i.e., words with opposite lexical meanings.

Competition and consensus between *fear* and *curiosity* give rise to a system for processing environmental information that enhances the adaptive potential of both individual organisms and the species as a whole. There is a close relationship for our brain between perception and actions. Our body serves us to perceive the outer world. We interact with it by the agency of our body and thereby correct brain function. Virtual models of the brain frequently tend to fail. Some faults can be useful, but others are dangerous, for two reasons.

On the one hand, they help us to understand the surrounding world. The faults allow dubious interpretations of any situation, e.g., when two different objects in the outer environment evoke similar sensations. Such problems are usually solved taking advantage of the fact that one of the possible interpretations is much more probable than the other.

On the other hand, the brain may err when a seemingly improbable interpretation proves correct. Such a situation gives rise to many optical illusions. A carrot irradiated by blue light is perceived by the brain as orange-colored, although all objective measurements of the reflected light wavelengths fall within the blue ($\lambda = 0.43-0.48$ μm) rather than orange ($\lambda = 0.60-0.66$ μm) range. The cause of such paradoxical perception lies in the virtual carrot model in the brain arising from our previous experience of seeing the carrot under different illumination.

To sum up, *our perception depends on a priori beliefs and past memories. The original perception of fear compensated by curiosity continuously improves our virtual model of the surrounding world. Neither vision, hearing, touching, nor other receptors are given preference in the correction of the model, because they all may serve as sources of information.*

These abilities are lacking in the computer systems of AI of robots having no body sensors. Robots devoid of a body equipped with receptors cannot learn on their own to forecast, nor can they correct the model by their actions.[3] Yet, there is more to it than that.

2.7 Sensors may also be wrong. Evolution of false image prevention

The retina of our eyes performs primary processing of light signals from environmental objects; it is frequently treated as a 'peripheral brain fragment'. Richard Feynman described the human eye structure as follows [53]: "... *the light-sensitive cells are located in the retina at depth, so that the light has to go through several layers of other cells before it gets to the receptors. Retina looks as if it is turned inside out! So some of the eye's structure features are wonderful, and some are apparently stupid*".

Describing an octopus' eye in the same lecture, he argued that: "*In the octopus it turned out, amazingly, that the retina is also a piece of the brain that has come out in the same way in its embryonic development as is true for vertebrates, but the interesting thing which is different is that the cells which are sensitive to light are not located behind the layers of other cells, as in our eye, but directly on the inner surface of an eyeball, and the cells which do the calculation are in back of them. So we see, at least, that there is no good reason for its being inside out. The other time Nature tried it, she got it straightened out*!"

However, Feynman's assertion is wrong. Octopuses are a phylogenetically older group than vertebrates. It is not that Nature had to correct its error when it created octopus eyes. Usually, Nature does not err; the species becomes extinct when it does. Deeply sunken human eyes are a remarkable adaptive trait. For sharp sight and reliable recognition, visual receptor cells (rods and cones) must be embedded in a medium having a practically constant or slow-changing temperature. Otherwise, rapid fluctuations of environmental temperature creating thermal gradients (e.g., caused by the wind) would change the rate of biochemical reactions involved in conversion of light quanta into electrical signals in the retina [54]. It would produce false visual illusions (*phantoms*) nonexistent in reality. The brain counteracting such a situation begins to behave chaotically in search of the answer to the questions 'What is this? Reality or illusion?' To obviate this problem, natural selection put the retina of terrestrial warm-blooded animals deep in the eyeball, where the temperature remains practically constant. Octopuses live in water with slowly changing temperature that has no appreciable effect on visual perception.

3. Adaptation of robots and humans to the environment

3.1 20th century: from cybernetics to synergetics

Over 70 years ago, the Nobel laureate in Physiology or Medicine 1973 Konrad Lorenz wrote that in the animal world *training practically always leads to adaptation, i.e., consensus with the environment*. In 1941, he published an article entitled "Kants Lehre vom Apriorischen im Lichte gegenwartige Biologie" [55]. Lorenz argued in his indirect dispute with Immanuel Kant that "*a priori* forms of thinking and intuition have to be understood as any other organic adaptation", i.e., as achieving a compromise that increases stability of the organism.

Rapid development of the sphere of application, high workspeed, and miniaturization of computers 30 years ago provoked the temptation to create human-like AI [56–62]. Sixteen years later, Herbert A Simon, another Nobel Prize winner in Economic Sciences 1978, stated in his book *Models of Man* [63] that within the following 10 years computers would be able to beat world chess champions[4] and write poetry and music. This prediction has come true. More than that, in 2016, a computer won a professional player of the game of Go (an ancient Chinese game offering a greater variety of combinations than chess).[5]

[2] The term 'antipode' appeared for the first time in Plato's dialogue *Timaeus* written c. 360 BC.

[3] Humans with an injured frontal lobe of the brain cortex frequently experience a similar difficulty. They lose the ability to correct the virtual model of the outer environment and, therefore, the purpose of their actions. Such people are unable to implement any plan or comply with instructions; they are always distracted by everything that comes into their view.

[4] The most inspiring success was the victory of the Deep Blue computer (IBM) over then World Chess Champion Garry Kasparov on May 11, 1997 by a score of 2 to 1 with three draws in a series of 6 games.

[5] In a series of five Go games between the Alpha Go computer program and the Korean professional player Lee Sedol (March 9–15, 2016), the computer won by a score of 4 to 1. The Alpha Go software was supported by 1,920 processors and used 280 graphics processor units operating in the distributed network. The games were televised live on YouTube.

Nevertheless, Simon allowed for some limitations. The lack of time needed to solve a problem necessitates consideration of only the most essential of the infinitely many possible approaches. Humans, unlike robots, try to elaborate a strategy (probably far from the optimal one) for a limited time, thus ensuring instantaneous success in a given situation [63, 64]. The key words here are 'limited time' and 'success'. A time deficit arises from the necessity to neutralize the influence of interfering and not infrequently life-threatening environmental factors. This thought was postulated by M M Bongard of the Institute of Biophysics (presently Institute of Theoretical and Experimental Biophysics, Russian Academy of Sciences) in the 1960s [60, 65].

In the 1950s, the Russian biophysicist M L Tsetlin initiated consideration of social matters from the standpoint of game theory. These studies had been preceded by those of J von Neumann [66], who analyzed situations in which the *players were a priori not fully informed about the game algorithm*. M L Tsetlin showed that robots using even a linear tactics could be taught how to cope with the task of adaptation to the game rules provided that the random behavior of the external environment changed only slowly. This suggested the crucial role of the factor of time in the formation of basic adaptive capacities. Based on this finding, Tsetlin turned to considering social games involving a large number of participants. The results of these studies were collected and summarized in a volume of his works [62], unfortunately published in 1969 after the author's death.

In 1990, the Australian roboticist Rodney Brooks published the article "Elephants do not play chess" [67], calling on developers of robots to create AI systems based on the results of neurophysiological studies. The author emphasized that the human brain is first and foremost a tool for adaptation to the environment. Positive emotions tend to stop panic attacks and promote problem-solving even under conditions of an information shortage by allowing one to go beyond the limits of the possible (catchphrase: *Nothing ventured, nothing gained*) [37].

Today, many researchers in the field of robotics (e.g., Moshe Vardi of Rice University, Houston, USA) [68] contend that there are no bounds whatever for CARs to become cleverer, quicker-witted, and more skilled than humans. However, this dream is difficult to turn into reality. I was once fascinated by the way the Hopfield neural network [69] operates, because it seemed to be capable of self-correcting balance problems [70]. Later on, I understood that the procedure of searching for equilibrium is determinate, because this ability was envisaged in the software and the algorithm was assigned by a programming specialist. Although the time needed for recovering equilibrium depends on a variety of external and internal factors (as exemplified by the model considered in Section 2.4), the procedure in the simplest case is nothing more than the search for the local minimum of energy (a 'potential well') on an n-dimensional cube.

All further modifications of this algorithm, e.g., the Hamming neural network for binary vector classification or different types of perceptrons [71–73], also lead to systems with a *rigorously determinate algorithm* developed by the programming staff. The 'thought process' in modern learning robots is only the next step in the development of *deterministic* Turing machines integrated into parallel-sequential networks [59]. Humans are *free to choose any goal*. The goal for a robot is formulated by the human, which means that robots lack inventive ability unless it is imparted by the developer.

The term *creativity* can be interpreted as the search for a spontaneous asymmetric response to an obstacle hampering the choice of the way to a set or an input goal, taking into consideration possible risks, failures, and pieces of luck.

3.2 Synergetics — theories of cooperative dynamic interactions between dissipative systems

In the 1990s, we used a population of colony-forming bacteria as a model of collective behavior. We intended, in the framework of then ongoing discussions, to ascend a set of stairs leading to the formation of *consciousness*, from the simplest to highly organized forms of life. Colony formation by bacteria was chosen as the starting stage of the self-organization process [74, 75]. The work was not completed for a lack of comprehensive knowledge about transition from the simplest to the most advanced organisms.

In the 1980s, synergetics [76] and the thermodynamics of dissipative structures developed by I R Prigogine and his school [77] replaced the cybernetics of the 1950s–1960s. The advent of synergetics, like that of cybernetics a few decades before, gave hope that new ideas would provide a basis for a *general theory* and give an answer to the question 'How do humans think?', bearing in mind that the human brain functions like a determinate Laplace system periodically passing from deterministic to chaotic behavior and back. The concept of mental process *chaotization* has a long history. In the late 1950s, Niels Bohr [78, pp. 27, 28] noted: "*The main traits of living organisms depend on the processes of atomic scale where we encounter important limitations on the application of classical physics concepts... . It follows from classical physics determinacy that any perturbation in a system consisting of a huge number of parts inevitably leads to chaos. In quantum physics, such description reflects the result of interaction between stable atomic systems; therefore, it is based on the outcome of competition between different individual processes. These processes determine in a simple way the state of new systems through atomic particles they contain as they would determine the initial state of the system...*".

In 1971, David Ruelle and Floris Takens published an article under the title "On the nature of turbulence" [79], based on modeling nonlinear dynamics. They criticized the Landau–Hopf scenario [80, 81] and provided evidence that dynamics may become turbulent after 3 or 4 bifurcations, and the system, in particular, may acquire a continuous spectrum characteristic of a random process. The authors attributed this fact to the emergence of a 'strange' attractor in the phase space due to trajectory instability. It was shown in experiments that many open nonlinear far-from-equilibrium systems experience self-organization [82] accompanied normally by originating either spatially nonuniform stationary (i.e., slowly changing with time) formations called dissipative structures by Prigogine [83] or periodic and quasiperiodic spatial–temporal waves referred to as autowaves [84, 85].

The International Union for Pure and Applied Biophysics (IUPAB) was set up in 1961. The agenda of IUPAB congresses required structurization of the totality of submitted reports. Therefore, it was decided to distinguish three disciplines, viz. *molecular biophysics, cellular biophysics, and biophysics of complex systems*. This classification was adopted and is still in use in biophysical journals, including the

Russian journal *Biofizika* (Biophysics), even if it is far from ideal, since it is *difficult to find simple systems* in biology. After all, what is complexity and what is the difference between simple and complex systems?

Over 30 years ago, a workshop on synergetics was held in Bavaria to discuss various aspects of complex systems, including the brain. However, there is thus far neither a universal theory of complexity assessment nor a universal approach to the evaluation of stability of biological systems because of their enormous diversity [27]. Their main property is defined by the formula '*changing anything changes everything*'. The human brain hosts numerous chain reactions, and a change in one activity triggers another [86]. Such a behavior is characteristic of *dynamical systems containing an essential chaotic component* [87].

G G Malinetskii and A B Potapov [88] attributed 'complexity' to the duality of information processing. On the one hand, it can be understood as complexity of a device, i.e., a system with a large number of elements and/or nontrivial links between them. On the other hand, it can be interpreted as the complexity of external manifestations of the system, irrespective of its internal structure, i.e., complexity apparent as a nontrivial behavior. The formation of dissipative structures and autowave processes is based on self-organization involving the presence of variable parameters of two kinds. Most of them correspond to 'soft rapidly rearranging regimes', while the remaining ones are regarded as belonging to 'hard modes'. After a rather long lapse of time, the former adjust themselves to fit the latter, i.e., the two systems of parameters merge into one. To be described, the resulting integrated system must attract new variable parameters not necessarily co-incident with the previous ones characterizing the behavior of two separate systems. In other words, a system that underwent self-organization needs to be described anew (see Section 2.4 for the simplest example). This property of self-organizing systems is called *emergent*, i.e., coming into being suddenly. In synergetics, this term is used to describe a system acquiring specific properties not inherent in its constituent components, taken either separately or collectively; these properties arise by virtue of backbone (system-forming) links between individual elements.

Today, the era of dissipative structures and autowave processes in synergetics is being succeeded by the era of self-organized criticality, drawing ideas from neurophysiology, biophysics, psychology, theoretical history (Big History), and other sciences analyzing the influence of *the past* on *the present* and *the future* [24].

Yu A Danilov noted in book [89] that the advent of the science of nonlinear interactions compromised Newton's concept of trajectory as a geometric line, i.e., 'a length without width'. Physically, the description of the behavior of a dynamical system in the language of trajectories would require an instrument with a resolving power high enough to 'see' a separate geometric line. Certainly, any real instrument has a finite resolving power, and the user can see a bunch of concurrent individual (sometimes entangled) trajectories instead of an individual line. The human eye is unable to distinguish between trajectories within the bunch; therefore, one can speak only of *a certain probabilistic distribution*. It is impossible to move from probability distribution to individual trajectories in the framework of Prigogine's and Stengers' terminology [90]: they are simply invisible to an external observer. Irreducible probabilistic distributions radically altered the description of dynamical systems and even the understanding of physical laws. The key words here are '*probabilistic distribution*' and the '*special role of the observer*'.

Prigogine and Stengers wrote in book [90]: "*Traditionally, there were two formulations of physical laws: one in terms of trajectories or wave functions, the other in terms of statistical ensembles. But such statistical formulation was not irreducible. It was fairly well applicable to individual trajectories or wave functions. In other words, statistical approach excluded the appearance of new dynamic properties. As a result, irreversible approach to equilibrium was traditionally associated with approximateness, 'coarseness' of description, and the appearance of the arrow of time with incompleteness of our knowledge. The proposed irreducible formulation breaks radically with the past. Irreversibility and probability become objective properties. They reflect the fact that the observable physical world cannot be reduced to individual trajectories and separate wave functions...*".

It follows from the foregoing that it is impossible to fully understand brain function *based on an analysis at one or two hierarchical levels*, e.g., a neural network or a sandwich-like set of networks. In such a case, the only informatively meaningful event is the exchange of electric impulses during movements of K^+, Na^+, Cl^-, and Ca^{2+} ions between individual nerve cells that trigger a cascade of further processes via synapses. The questions concerning the human brain that remain open are: "How is it organized to be able to simultaneously perform determinate and chaotic functions?" and "How can physical macrosystems exist based on duality?"

Any human is always something more than an observer of his/her behavior can directly perceive. Human intelligence viewed as a whole does not obey the law of identity: $\{A_{\rm obs}\} \neq \{A_{\rm real}\}$, where $\{A_{\rm obs}\}$ is the observed set of intelligence characteristics, and $\{A_{\rm real}\}$ is the real set of its properties. Intrinsic in each human are hidden parameters that manifest themselves upon a change in the surrounding conditions, but otherwise remain *a priori* unapparent; hence, the inequality

$$\{A_{\rm real}\} > \{A_{\rm obs}\} \tag{20}$$

is satisfied.

In other words, human intelligence and its behavior are basically multivariant. A human does not fully open themselves up to an external observer.

Until quite recently, robots had been considered to be merely machines, i.e., something well-defined, but the behavior of certain modern self-learning robots resembles human behavior. It is sometimes difficult to deduce from a computer processing self-learning AI information the indicators it utilized in solving a given problem.

Moreover, both the robot and the human are open systems for the inflow of energy and environmental information. In the human, the openness means the inability to fully understand the rules that govern their brainwork, consciousness, and behavior. These rules vary continuously under the influence of both the outer and inner environments. Suppose that a person is perfectly aware of the algorithm $\aleph$ of their own behavior comprehensively describing the functioning of their psyche and brain $\{A_{\rm real}\}$. Such a person would be able to accurately predict risks associated with decision-making and act as appropriate to avoid their unforeseen consequences.

Suppose further that this knowledge enables them to conclude that the action P in situation X will be fraught with troublesome consequences. It will then be possible to avoid them by consciously refraining from action P and acting otherwise.

However, a decision is always taken when the behavior of a competitor in the outer environment is uncertain. For this reason, *risk assessment is always probabilistic*. In other words, humans as a deterministic system must perform action P and must not do it as a self-aware probabilistic system. The uncertainty leads to an oscillatory process; hence, Hamlet's eternal question to be or not to be? To act or not to act?

4. 21st century: new interpretation of old ideas

4.1 Time of expensive projects

The studies by Henry Markram [91–107] allowed him to put forward in 2009 the idea of a supercomputer-based mathematical brain model [108]. The idea seemed simple: the author believed that a software program simulating the entire network of 86 billion neurons and 100 trillion synapses corresponding to human brain architectonics would ensure a transition from quantity to quality, i.e., a novel system reproducing all cognitive properties of the human brain. Markram's idea provided to be the basis for the 10-year (2013–2023) Human Brain Project worth a total of \$1.63 billion.

The expediency of the project was questioned within a year after its proposal [109]. More than 800 neurophysiologists wrote letters criticizing the project. Its opponents' arguments reduced to the fact that the work of the human brain in comparison with that of computer–based AI remains poorly understood. A commission was set up that summarized arguments against Markram's project in a 53-page report. The critique was extended to include inter alia the decision-making policy and practice of Brussels concerning expensive research projects approved without their scientifically sound analysis.

In the context of Markram's project criticism, a \$30-mln megaproject (Human Connectome Project) initiated in the USA in 2009 [110] should be mentioned. Here, the term 'connectome' implies a connection with an Internet access server. Connectome is a dynamic graph (map) of the links between self-organizing subsystems of the brain that arise, develop, and decay under the effect of long-range forces and fields as exemplified by axonal guidance and pathfinding, i.e., the directional growth of axonal and dendritic networks. In the developing nervous system (during ontogenesis), axons extend over long distances to reach all cells and organs and thereby form architectonics of neural connections as an integral system of the body. In this regard, the body and the brain make up a whole. The study of forces and interactions influencing the pathfinding is of importance for understanding the mechanisms that underlie mind and body unification, as well as faults in their self-organization resulting in pathological changes.

It was believed that the mechanisms behind the formation of neural links are responsible for the main aspects of human individuality, such as personality, intelligence, and creativity. Therefore, modeling the human connectome could become an important step toward understanding all the variety of mental processes provided that an adequate programming language is created [111]. In Russia, an adherent of this idea is K V Anokhin who uses the term *cognitome* (brain's hypernetwork) [112].

In August 2017, Chinese researchers announced the opening of the Suzhou Institute for Brainsmatics, affiliated with the Huazhong University of Science and Technology (HUST), which will be focused on brain mapping. Four hundred fifty mln yuan (67 mln US dollars) were allocated from the state budget to the new research center for a period of 5 years to support brain mapping studies. Their main objective is to improve existing brain mapping technologies by applying electron microscopy of thin nano-scale tissue layers for further investigations.

The Allen Human Brain Reference Atlas created at the Allen Institute for Brain Science, USA, has recently become available online. Its printed version published by a group of 40 authors occupied the whole of the *Journal of Comparative Neurology* volume issued on 15 September 2016 [113].

However, my personal experience [114, 115] gives evidence that investigations into brain anatomy are of little value for understanding the mechanisms of creativity, unless they are supported by simultaneous studies of brain function dynamics. Each person's brain is a unique and individualized organ that keeps track of the life experience of its host. The probability of complete coincidence, even in identical twin siblings, is close to zero. Moreover, any idea requires experimental verification. To study the dynamics of changing links, one needs a method to observe the formation of both the structure and the function of the same brain. The processes in neural networks and their interstitial space proceed in millisecond time ranges and on nanometer spatial scales. Unfortunately, the resolving power of methods currently available is far from such temporal–spatial resolution. The temporal resolution of positron emission tomography (PET) is too low (~ 10 s) [116]. Electroencephalography (EEG) has a temporal resolution up to 2 μs [117] but low spatial resolution. The implantation of microelectrodes increases spatial resolution, but this invasive method cannot be used in routine human studies and is applied only in exceptional circumstances for therapeutic purposes. Moreover, all invasive techniques pose an uncertainty problem, because introducing a foreign body into the brain causes adverse reactions. MRI can be used to record changes in oxygen consumption and therefore in metabolic rate. The spatial and temporal resolutions of MRI are around 1 mm and 2 s, respectively [118].

High hopes are placed on advances in neurophotonics, e.g., for the spatial–temporal analysis of the function of neural networks with the aid of flexible fiber-optic devices to simultaneously excite and register inherent or induced fluorescence of neurons. The temporal resolution of this method is formally unlimited, while its spatial resolution depends on the wavelength of radiation being used [119]. This method is actually a miniaturized variant of dynamic nano-endoscopy applied to brain research. However, neurophotonic techniques are invasive methods, like the implantation of microelectrodes, sharing the disadvantages and uncertainties of that approach.

A book by the Sheroziyas (father and son) [120] also presents a plan for the construction of an anthropomorphous robot. The estimated cost of the 10-year project amounts to \$10 billion. The authors take into consideration feedback between the brain and the body. According to them [120], the mean number of degrees of freedom in a moving human body is on the order of 300, although I for one doubt this estimate. The human organism can be regarded as a condensate of unicellular organisms formed during evolution in a single body volume [27]. This giant colony is governed at all levels by a universal genetic code. Such a dynamical system must have roughly as many degrees of freedom as the number of cells in the body of an adult human, i.e., $\approx 10^{13}$ [27].

4.2 Classical and probabilistic logics

Let us try to describe the mind and consciousness in a different language, i.e., proceeding from the logic of the functioning of brain networks to their architectonics, rather than vice versa. There is nothing new in this approach to the analysis of dynamic processes: it was used by many researchers even before the 20th century and continues to be applied in the 21st century. Suffice it to recall the lecture delivered by V G Red'ko at the IV All-Russian Scientific and Practical Conference 'Neuroinformatics-2002' [121]. It will be shown below that this approach yields a nontrivial result.

Classical logic, unlike probabilistic logic, uses well-defined rules. The former provided a basis for the mathematical apparatus of classical physics. It includes all aspects of logic represented in symbolic form. An important contribution to symbolic formalization of the logical thinking processes was made by O de Morgan [122], G Boole [123], F L G Peano [124], G Frege [125], A Poincaré [126], D Gilbert [127], B Russel [128], and other scientists.

However, the potential of classical logic for obtaining objective knowledge has its restrictions. Classical logic unambiguously identifies environmental objects with their symbols (the law of identity) [129]. An algorithm should be designed in a consistent way following a linear cause $\rightarrow$ effect chain. The logical law of contradiction forbids self-contradiction in reasoning and analyzing situations. For example, contradictory statements cannot both be true in the same sense at the same time. Moreover, the law of excluded middle prohibits answering questions indeterminately: neither a yes nor no. Finally, a thought is called true only when it ensues from another thought on which it is based (the law of sufficient reason). In other words, thinking must be a consistent process.

Humans feel comfortable in the framework of this coherent system of logical thinking, because classical logic is closed, internally consistent, and unambiguous. However, it follows from practical experience that a closed system of assertions contains statements that are impossible to categorize as either true or false. The most striking examples are paradoxes, e.g., Zeno's aporias (ancient Greek *απορια* — difficulty, perplexity) [130], that cannot be explained in terms of classical deterministic logic.

The boundedness of classical logic is also due to the fact that it reflects a single class of environmental situations corresponding to 'order' as opposed to 'chaos' [131]. It is difficult to explain, in the framework of classical logic, the spontaneous formation of a goal by a human or to predict the trajectory of development of nonlinear processes, such as freaks of weather, earthquakes, changes in economic, biological, and social systems, or manifestations of creativity. Attempts to build the future out of the past in nonlinear systems or formulate political doctrines based on formal logical considerations not infrequently lead to dogmatism and, as a result, to the collapse of closed worldview concepts.

It has been widely thought since the time of Parmenides [6] that "*only cognizable beingness rather than sensory impressions is a real-world entity*" [132]. Parmenides' hypothesis can be interpreted in a modern light based on the following postulate: *laws* (*models*) *of nature must not be more complex than the data they explain*. Otherwise, the notion of information compression is invalid, since any unprocessed set of data can exist by itself as a 'law'. If memory were not limited, it seemingly would be possible to remember all situations without distinction and treat them as particular natural laws. However, real laws of nature formulated by science are interpreted based on a different principle — *the principle of simplicity, memory saving, and experimental verification* [133].

At the end of the 20th century, Gregory J Chaitin showed in book [134] that this line of reasoning leads to contradictions. The overwhelming majority of the lines of symbols, e.g., sequences of random numbers in irrational fractions, describing environmental phenomena appear to be incompressible in principle. If so, they cannot be reduced to simpler, shorter, sequences. Such a situation takes place whenever an observed chain of events exhibits no internal patterns allowing it to be compressed. Then, there is no choice but to accept the entire sequence as the law of nature. We accept this situation without assigning any reason merely because we have no other choice.

It was noted above that the brain can design, all on its own, virtual models (lines of programs) of any degree of complexity. Their comparison with reality makes it possible to reveal local (special) patterns existing in the outer environment. A dynamic comparison is realized by comparing a synthesized sequence with that conditioned by the environment, which exposes local patterns separated by large intervals in time and space in the case of fragment coincidence, e.g., words and phrases in the text. In 1989, we realized such an idea in the form of an algorithm. Spectral analysis in the Fourier space made possible the search for repetitions separated by large intervals, e.g., in nucleotide sequences of DNA [135].

Finding repetitions allowed us to compress information, with occasional incompressible parts of a sequence being regarded as 'trash'. However, it was impossible to explain why the 'trash portions' did not contain information. A change in the starting fragments changed the 'trash' composition. The sole explanation for such a situation is that the compression was performed in the framework of classical deterministic logic, does not contradict it, and permits concrete practical problems to be solved: here a trash serves as a noise. The actions of robots are governed by algorithms based on the laws of classical logic assigned by a programmer. Brain work consists of overcoming contradictions between antipodes, i.e., between the present and the past, environmental events perceived by sensory organs and memories of them conserved in previous experience.

Dialectical logic is often described as arising from competition in disputes (the saying: *truth springs from argument*). B B Kadomtsev, with whom I several times discussed the emergence of *the probabilistic world of information* technologies, insisted on its relation to quantum mechanics [136, 137]. He summarized his views in the book entitled *Dynamics and Information* [138].

[6] The Greek philosopher Parmenides of Elea (late 6th–mid-5th century BC) tried to distinguish between the truth and opinions about it. Zeno, the author of famous paradoxes (aporias), was his disciple and successor.

In the early 20th century, David Hilbert called on mathematicians to define the finite totality of principles allowing, based on the consistent application of the rules of classical mathematical logic, us to find harmony in mathematics. By 1922, Hilbert elaborated a plan to make mathematics closed by means of its complete formalization based on the proof of the absence of inner contradictions. To implement this program, Hilbert developed, in a continuation of Frege's work [125], a logical theory of proofs allowing the principle of consistency of mathematics to be reduced to the proof of consistency of, say, arithmetic. Hilbert used for this purpose only universally accepted means of classical, rather than probabilistic, logic. G Chaitin arrived at the conclusion [134, 139] that Hilbert's program proved infeasible, even if it gave a powerful impetus to the development of logic.

In his time, Kurt Friedrich Gödel was also interested in a similar problem. He showed in 1931 that formalization of mathematical theories encounters the inconsistency problem. 'Improvable theorems' appear at the boundaries of any closed sets [140]. Gödel argued that any theory is incomplete and therefore contradictory. The incompleteness implies the presence of statements that *can be neither proved nor disproved from axioms of a given theory*. Inconsistency leads to the appearance of *paradoxes, i.e., the possibility of proving any assertion by considering true as false and vice versa*.

The history of science gives evidence that everything new is actually well-forgotten old. In 1686, i.e., 250 years before Gödel, Gottfried Wilhelm von Leibniz raised in his *Discourse on Metaphysics* [141] the question of how to distinguish facts described by a certain law from those indescribable by any law. Leibniz re-formulated the thought of Parmenides and came up with a simple postulate: *a theory has to be simpler than the data it explains. Otherwise, it is of no value*. What does the word 'simpler' mean here? In modern language, it means that information must be compressed and written down more briefly than the original data. But such a compression should be performed following a rule that allows expanding compressed information, if appropriate (see Section 5).

The human brain compresses information. The number of gradations at hierarchical levels for the formation of behavioral algorithms in living systems remains to be determined but appears to be very high, i.e., comparable to the number of organelles or cells ($10^{13}-10^{17}$).

Thus, the surrounding world and the human as its integral component are organized hierarchically in accordance with the laws of both classical binary logic and mixed ternary logic supplemented by probabilistic logic. As a result, the brain functions in the ternary system: 'yes–no–uncertainty with a different degree of probability'. A brain living in its own virtual world divorced from reality is the brain of a madman. Models it synthesizes may sometimes be close to reality, but this looks as the exception rather than the rule.

4.3 The brain living in the probabilistic world

Let us call a brain living in the probabilistic world a Bayesian brain. The simplest formula taking account of conditional probability (the Bayes formula) was known a long time ago [142, 143]. Consideration of conditional probability $p(\mathrm{A}|\mathrm{X})$ expands our knowledge about event A upon obtaining new data about event X:

$$p(\mathrm{A}|\mathrm{X}) = \frac{p(\mathrm{X}|\mathrm{A})p(\mathrm{A})}{p(\mathrm{X})}\,. \tag{21}$$

Probability $p(\mathrm{A})$ is the belief-based probability of a certain event A, i.e., fiducial (Latin fides — belief, confidence) probability. Probability $p(\mathrm{X})$ is the probability of the appearance of new information that can change our *a priori* belief in the realizability of event A. Conditional probability $p(\mathrm{A}|\mathrm{X})$ makes it possible to estimate to what extent our conviction in the probability $p(\mathrm{A})$ of event A was justified. In the case of absolute confidence in its reality, the probability $p(\mathrm{A}) = 1$. In the case of absolute confidence that event A will never occur, $p(\mathrm{A})$ tends to 0. Most situations are intermediate between 0 and 1. The nearness to 0 or 1 changes as new data become available with probability $p(\mathrm{X})$.

The popularity of the Bayes formula seems to be ascribed to the possibility of correctly calculating the probability of attaining a goal after obtaining new information. It forms the basis for the concept of the so-called *ideal Bayesian observer*, an imaginary person always using the available data in the best of all possible ways [144]. This imaginary creature, however, resembles another one known in thermodynamics as Maxwell's demon [9]. To recall, the Bayesian observer and Maxwell's demon are connected via the *cost of action* (the function of energy and time; see expression (2)). They cannot measure at no charge and accurately the magnitude and direction of a probability change upon receiving new information. But the ideal Bayesian observer, unlike Maxwell's demon, *'feeds' not only on energy but also on time spent on repetition* (saying: *repetition is the mother of learning*).

When using conditional probabilities $p(\mathrm{X}|\mathrm{A})$, the following circumstances should be taken into consideration in the Bayes formula for three reasons.

First, a large number of trials are needed to reliably deduce the probability of a certain event A. Only in this case does probability $p(\mathrm{A})$ allow an objective estimate of the possibility of the occurrence of this event.

Second, one must be sure that probabilities of events A and X are interdependent. If neither $p(\mathrm{A}|\mathrm{X})$ nor $p(\mathrm{X}|\mathrm{A})$ changes the probability of event $p(\mathrm{A})$, then A and X are independent.

Third, in the general case, $p(\mathrm{A}|\mathrm{X}) \neq p(\mathrm{X}|\mathrm{A})$. A chain of conditional probabilities can be asymmetric. In the case of event A, event X occurs with probability $p(\mathrm{A}|\mathrm{X})$; the opposite assertion that in the case of event X event A occurs with the same probability may be wrong. *The process of human thinking tends to ignore general information about the frequency of events and focus on special information about the event interesting for it alone*; *therefore, it misrepresents the real situation described by probabilities*. For example, the prevalence of cancer in the general human population being rather low, the confidence that an individual person does not have cancer is close to unity, i.e., $p(\mathrm{A}) \approx 1$. If a person has a growing tumor (event X), the probability of a positive result in the cancer test is $p(\mathrm{A}) \approx 1 - p(\mathrm{A}|\mathrm{X})$, the cause being the rarity of malignant tumors compared with the large number of benign ones. The probability of this event is about 15%, i.e., $p(\mathrm{A}|\mathrm{X}) \leqslant 15\%$. The confidence that a person does not have cancer also changes but insignificantly, i.e., probability $p(\mathrm{A}) \approx 85\%$. Finally, in the case of a positive result of a cancer diagnostic test using a modern method with an accuracy of 90%, probability $p(\mathrm{A})$ for such person is close to 90% (saying: *while there's life, there's hope*).

For an external observer, e.g., a physician, $p(\mathrm{A})$, the probability of diagnosing cancer in a single patient among many others seen during routine prophylactic examinations (X) is very low, $p(\mathrm{X}|\mathrm{A}) \leqslant 1\%$. In other words, matching two conditional probabilities is fraught with errors caused by *base rate* neglect in setting up the problem of the initial frequency of a given event. The base rate notion being of great importance, and here is one more example to clarify it in more detail.

A topical example is the new technology to automatically detect terrorists (base rate fallacy on https://ru.wikipedia). Suppose a million-plus city happened to become a sanctuary to 100 terrorists per 999,900 law-abiding citizens. The police installed an alert system with cameras to automatically identify criminals' faces in public areas based on the relevant database. The software is likely to make errors of two kinds: to overlook a sought-after individual (terrorist) with a probability of 1%, and to mistake a peaceful citizen for a terrorist they resemble (false alarm) with the same 1% probability. When a camera sees a terrorist, the probability of mistaking him for an ordinary citizen is 1%, meaning that the correct signal informing police about the appearance of a terrorist will be sent in 99% of cases and will be absent in 1% of them. When the camera sees an ordinary citizen who looks like a terrorist, it produces a false alarm signal in 1% of the cases.

What is the probability that a person who causes the camera to send an alarm signal is a terrorist?

Setting aside the above-mentioned three conditions, it can be assumed that the probability equals 99%. Although this assumption seems correct, it is actually wrong: the true probability is around 1%. The discrepancy is due to the confusion of two values of different natures. *The frequency of the absence of signals per 100 terrorists is unrelated to the number of signals for every 100 peaceful citizens.* The error is easy to understand by considering a limiting case in which the face-recognition system operates in a city *free from terrorists.* Wrong signals are sent once for each 100 citizens, i.e., all of them are false. If, however, nobody is aware of it, the police have to arrive in response to each false signal, i.e., in 100% of the cases. Imagine now that all 10^6 residents of the former city walk past the camera. It will send signals about ~ 99 of the 100 terrorists, and 9,999 of the 999,900 ordinary citizens, because they constitute the majority. On the whole, the alarm signals will be produced when roughly 10,098 people pass by the camera, of whom only ~ 99 are terrorists. Therefore, the conclusion that the person who caused the camera to send a false signal is a terrorist is correct in 99 cases out of 10,098, i.e., in less than 1% of the cases, which is much lower than the initially assumed value of 99%.

Rare events frequently lead to wrong conclusions when analyzed using the Bayes formula. Jeremy Wolfe and co-workers from Boston undertook an experimental analysis of security services scanning passengers' luggage in an airport in search of knives, explosives, and other items prohibited for air transportation on board aircraft. The results of the search were in agreement with the Bayes formula, when the sought-after objects occurred frequently. The security agents missed only 7% of such objects. The result proved disappointing when such objects were rarely found. More than 50% of the banned objects were missed in only 1% of the examined luggage (see book [144]).

The Bayes formula becomes operable only after repetitions and data collection, i.e., following correction of the virtual environment model in the brain or a computer taking advantage of past experience gained in interactions with the real world. Otherwise, we find ourselves in the trap of false correlations [145]. In the past, the irresistible desire of the human to fulfil the lifelong ambition of finding order in the outer environment, even where it is intrinsically absent, was a source of superstitions [146]; at present, it gives rise to hypotheses whose authors mistake the wish for the reality [147].

The perception of the real world in the brain begins from an *a priori* belief that is actually a virtual model of the world in which objects and links between them occupy 'a certain' position in space and time. The model is determined by genetic factors and past experience. The brain utilizes this model to predict what signals must enter our eyes, ears, and other sensor systems. The brain, however, may be wrong. A comparison of these predictions with real signals reveals errors that prompt the brain how to improve its own environmental model. The cycle repeats again and again until the errors become negligible. The number of such cycles is usually small (in fact, only two of them are needed in the limit). It usually takes the brain a minimum of 100 ms to perform them (the limiting time of a simple motor reaction). *Our body serves to perceive and cognize the world.* Perception is a cycle in which predictions are continuously checked by testing against actions. This ability was absent in early AI and CAR systems.

To recall, the brain acquires a new image of the external world after updating and can repeat the procedure based on a new prediction of the character of events perceived by sensory organs. Each repetition of the cycle reduces mistakes. As soon as they become small enough, the brain can orient itself in the outer environment. It is easy to determine how quickly probability $p(\mathrm{A}|\mathrm{X})$ increases with repetition of corrections.

Let us turn back to the above example of cancer diagnostics. If a patient once again undergoes the cancer diagnostic test, the uncertainty may be radically reduced. If the result of the second test is again positive, the probability of cancer $p_2(\mathrm{A})$, in accordance with the Bayes theorem, will be the same as in the first test, i.e., 0.5. In other words, the reliability of the diagnostics remained unaltered ($p(\mathrm{X}|\mathrm{A}) = 0.99$). The probability of positive and negative results of the test is equal, i.e., $p_2(\mathrm{X}) = 0.5$. Substituting these values into expression (21) yields

$$p(\mathrm{A}|\mathrm{X}) = \frac{0.99 \times 0.5}{0.5} = 0.99\,. \tag{22}$$

It means that the probability of cancer in a given patient is 99% instead of the former 50%. This example shows that repetition quickly unravels the tangle of false and true situations and makes the answer more precise.

As far as repetitions are concerned, the special role of a teacher should be emphasized. The 'teacher' reduces the proportion of errors [61]. Certainly, communication with the environment is not the sole way to gain experience. AI programs seem to be free from this disadvantage, because the programming personnel (the teacher) insert their own experience into them. However, a robot tends to act like a bull in a china shop, if the algorithm of the program is designed to obtain advantages in a given situation in the absence of rules for avoiding losses. Since the robot is unfamiliar with such notions as ethics, humanism, and morality, it is likely to physically destroy everything that hinders it from achieving the set goal, meaning that the programming specialist must envisage in the mechanism of 'teaching' not only the shortest path to the goal but also the cost that should not be exceeded on the way to it. As soon as the boundary is reached, the Gödel theorem comes into action, which results in situations in which the true and false (opposite) goals are intermixed and indistinguishable; in other words, the stability of the problem formulated by the programmer is compromised.

The program rapidly and accurately performs that and only that which is put into it and not what the operator might have expected to obtain after it is accomplished. Norbert Wiener arrived at the same conclusion in the mid-20th century, which he outlined in the complementary chapters of his book [148]. Uncertainty is fraught with 'surprises' likely to be brought by AI to the developer of the program.

5. Conclusions

5.1 Mathematical compression of information

A hypothesis of the existence of superalgorithms yet to be found has been proposed that extends the metaphor of development of logic and mathematics based on the following postulate: *coding allows information content to be compressed*. For example, one should not necessarily remember an entire infinite sequence of natural numbers, arithmetic or geometric progressions, Fibonacci numbers, and many other convergent series. Suffice it to remember the formula for calculating any Nth number of the series needed to construct such sequences of any length. This means that they contain little information.

However, there can be certain infinite series of seemingly nonrepeating sequences of numbers with local peculiarities. Let us consider the calculation of the sequence of the number $\pi = 3.14159\ldots$.

There are many algorithms for its consecutive calculation, viz. the stochastic Monte Carlo method, Poisson integral, François Viète formula, etc. As far back as the 16th century, F Viète derived a formula for finding the infinite sequence of digits of this series [149]:

$$\frac{2}{\pi} = \sqrt{\frac{1}{2}} \times \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}}} \times \sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2} + \frac{1}{2}\sqrt{\frac{1}{2}\times}}}\ldots. \quad (23)$$

This expression implies that the calculation of π is a recurrent procedure. Evidently, the time of calculation in this way increases with increasing the number N of a given item of the series.

In 1997, David Bailey, Peter Borwein, and Simon Plouffe arrived at a different, more elegant, and less time-consuming formula for calculating a Nth digit in the π sequence based on the idea of local convergence of the pieces of certain series [150]. This formula allows calculating any Nth digit of the number π without computation of the immediately preceding and following members of the series:

$$\pi = \sum_{0}^{\infty} 16^{-k}\left(\frac{4}{8k+1} - \frac{2}{8k+4} - \frac{1}{8k+5} - \frac{1}{8k+6}\right). \quad (24)$$

To calculate the Nth digit of the number π using this formula, both its parts should be multiplied by 16^N in order to convert the factor in front of the parentheses on the right-hand side of formula (24) into 16^{N-k}. Then, the sum of several adjacent members of the series is calculated. There is no need to calculate many of these members, since the formula suggests that 16^{N-k} rapidly diminishes as k increases, so that subsequent digits have no strong influence on the value of the Nth digit in the series being sought. The BBP formula was derived using the PSLQ algorithm.[7] The history of mathematics records many other tools for information compression. Suffice it to mention Stirling's (or Moivre–Stirling) formula for computing the factorial and gamma function [151].

All this provides a basis for the hypothesis of a hierarchical level above a set of number series, i.e., a variety of algorithms for a series information compression. It can be assumed that another mathematical world exists above these algorithms. Its name, e.g., a set of superalgorithms or meta-algorithms, is immaterial. The superalgorithms of this level make possible the analysis of information about algorithms. In such *hierarchical variants*, mathematics is not a closed system: it is boundless, which opens up prospects for the creative pursuit of information compression modalities by moving from one hierarchical level to another. The higher the level, the smaller time needed to solve problems at the lower levels. Continuous extension of boundaries removes the limitations formulated by Gödel.

Two conclusions of importance for biophysics follow from this metaphor.

First, *when ascending stepwise the hierarchy scale, a way out must be found from the previous level to the next one leading into a different information space. In addition, the ascension from one level to another must be successive, i.e., not a single level can be missed.*

Second, *the algorithm must be insensitive to the initial conditions* [27], because the past (starting conditions) combined with boundary conditions (i.e., the present) defines the future.

Nevertheless, different kinds of phase transitions violate the laws of evolution, because they originate from jump-like transformations. The jumps can be beneficial at low levels as sources of *enlightenment*, creative activity, and advancement, but they are deleterious at the upper levels, leading to the downfall of the system or part of it; in the biosphere, they result in the replacement of one species by another. An advantage of the virtual world of our brain lies in the fact that it can reduce the probability of jumps (saying: *measure twice, cut once*).

5.2 Overrunning usual space into a new information space

The human brain solving intelligence problems tries to find a nontrivial approach, as exemplified by *the parable of the smart kid*. A father wanted his child to play in some way to be free to do his own work. He tore a page showing a world map out of a magazine, cut it into pieces, and said: "I shall take you to the zoo if you assemble the map". He was sure it would take the child at least an hour and a half, during which time he intended to complete his work. However, the son spent only ten minutes to cope with the task. "How did you manage it? Are you so good at geography?" asked the surprised father. "Daddy, you did not see the drawing of a man on the other side. I assembled the drawing and here is the map. Let's go to the zoo!" (borrowed from manual [152]).

[7] The PSLQ algorithm of analysis in the language of discrete questions is one of the variants of the search for the nonintegral relationships between the set of items of the sum of real numbers x_i with coefficients a_i. This sum is set to zero. The computation program either finds the integer x_i/a_i ratio or shows its absence (see PSLQ algorithm on the Wolfram MathWorld website built by Eric W Weisstein).

The child solved the difficult problem by moving into the information space familiar to him from *prior experience*. In other words, he compressed information and thereby saved time needed to solve the new problem.

The approach to determining π described in Section 5.1 shows that the time needed to compress algorithmic information depends on the level of the programing language chosen to address the problem. For example, algorithm operating time $\Delta\tau$ necessary to calculate any Nth digit of π is proportional to the serial number of the sought N digit, whereas program memory is proportional to $\ln N$ [150]. Because

$$N \gg \ln N, \qquad (25)$$

here is a simple but important conclusion: *memory and experience (knowledge of the past) save much time to solve new problems*.

The possibility of compressing information for constructing new models of environmental behavior depends on the possibility of integrating series or matrices having different properties. A review of the theory of matrix models from the standpoint of its relation to integrable hierarchies can be found in Ref. [153]. The mathematical foundation for the notion of complete integrability of dynamical systems, e.g., for the purposes of quantum mechanics, was laid by L D Faddeev [154].

However, divergent series pose algorithmic problems of information compression for a finite time interval. In the mid-20th century, A N Kolmogorov proposed using a program length measured in bits or computation algorithm l (which is the same thing) that transforms a given sequence $\{Y_{\mathrm{j}}\}$ into $\{X_{\mathrm{i}}\}$. Such a transition allows the information compression problem to be addressed in the general form [155, 156]. For a low-complexity problem (making compression possible), l is significantly smaller than the length of $\{X_{\mathrm{i}}\}$. In the opposite case, the process is algorithmically incompressible. For such sequences, the $l \sim N$ algorithm of the $\{X_{\mathrm{i}}\} \to \{Y_{\mathrm{j}}\}$ transition reduces to remembering consecutive symbols of the entire $\{Y_{\mathrm{j}}\}$ sequence.

The question is: are there absolutely random incompressible processes in nature, or are they mere mathematical abstractions? Possibly, the level of algorithmic hierarchy of humankind is not sufficiently high to answer this question. This problem interested Albert Einstein in the early 20th century (hence his well-known saying in the debate with Niels Bohr: 'God does not play dice with the universe' [157]). Bohr gracefully waived the discussion by proposing *the principle of complementarity* [158].

5.3 Limits to similarities between robots and humans

It is impossible to answer the question: Where's the limit of similarity between the robot and the human?. The search for the answer is the realm of science-fiction writers and futurists rather than physicists and biophysicists. It is impossible to separate facts from fables in *terra incognita*. Predictions of the future encounter intrinsic uncertainties. Nonetheless, we shall try to formulate conditions for the appearance of *a universal creative robot*. It was noted in a preceding section that creativity is the ability to set a goal and take creative decisions to achieve it. Modern robots cannot set goals by themselves. The goal is formulated by the developer who includes it in the AI program.

Purpose-oriented human behavior appears to have developed as a result of natural selection of the adaptive mechanisms designed to ensure survival of the species in the ever-varying environment on our planet. It is believed that hominids appeared 4.5 mln years ago and evolved over the following 2.8 mln years. A crucial event was the appearance of *Homo habilis* in Africa (the Lower Paleolithic). This was followed by a qualitative transition in human evolution and the beginning of social life. People gradually spread over the globe. The evolution of *Homo sapiens* side by side with other hominid species took approximately 3.5 mln years [159–161] (Fig. 11).

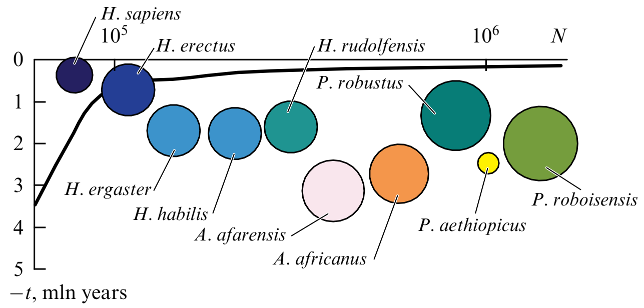

Figure 11. Stages of human evolution (t is the time elapsed, N is the population number) that ended with the origin of modern humans as a result of mutations in hominids that facilitated adaptation to the varying environment. Circle diameter corresponds to the population size; circle position is the lifetime of the respective species. *Homo sapiens* appeared some 200,000 years ago. It possessed the best adaptive abilities and formed the species that occupied virtually all regions of the planet. Now, the world's human population exceeds 7 billion people [162, 163].

What changes in the brain made us different from our ancestors? Anthropological findings give evidence that the brain of *Australopithecus* had a volume of 459 cm^3 on the average, comparable to that of certain chimpanzees. It increased to 930 cm^3 in *Homo erectus* (1.6 mln years ago) and reached 1330 cm^3 in *Homo sapiens* (200,000 years ago). It is not only and not so much the increased brain volume and changes in the brain structure or peculiar features of its frontal cortex (Broadmann areas 10) that participate in the construction of the virtual environment model (as demonstrated by EEG studies) based on genetically determined instincts and improve themselves, taking advantage of information coming from receptors (even though the frontal cortex of *Homo sapiens* is more than twice as large as in other primates). In all probability, the key factor is the 50% increase in the distance between neurons, which provided additional space for the growth of dendritic spines believed to be involved in memory formation and which contribute to the improvement in the virtual model of outer environment. Taken together, these changes laid the foundation for the appearance and development of the social communication language. This substantially promoted the formation of the virtual environment model and opened up prospects for the use of informal logic and creative activity by comparing and integrating information received by the brain cortex not only from sensor systems but also from the joint actions of different groups of people exchanging information and experiences [164, 165].

Brain development resulted in a survival strategy based on such individualized human qualities as fear of death, pain avoidance, the ability to evaluate potential reproductive hazards, offspring protection, food procurement, and the search for comfortable living conditions for oneself and one's relatives. This strategy was supplemented by cooperative behavior in response to environmental changes, remembering them, and extensive exchange of experiences.

Is it possible to integrate the virtual model of the surrounding world in the form of an algorithm into the AI of a self-learning robot? It is thus far impossible.

To begin with, we do not know such a model in full measure. The process of biological evolution remains to be simulated and reproduced in experiment. It is an incorrect inverse physical problem. Discussion of the origin of living matter would be purely hypothetical after such matter came into being. A peculiar feature of incorrect problems is the fact that they imply restoration of a process dating far back into the past based on present-day facts and events. Such problems are highly sensitive to initial conditions, which we know only in theory [166–168]. Thus far, we are aware (and then only partly) of a single variant of evolution, i.e., the development of living matter on Earth. It was shown above that the significance of a single observation calculated using the Bayes formula does not exceed 50%. This accounts for the enormous controversy regarding panspermia (extraterrestrial origin of life) [169, 170].

Second, robots have neither the genetic past nor cause-and-effect relations with it. Suppose, however, that humans created:

(1) a robot with a quantum computer-based 'brain' operating at the speed of light and a body packed full with a variety of receptors;

(2) a closed-circuit operation system (functioning without people) of self-reproducing robots (J von Neumann discussed the theoretical possibility of such devices in the 1960s [171]).

What comes next? The robots will further develop in the absence of humans provided there are sources of energy and raw materials available to them. How many years would it take their brains to go all the way (making use of the human experience passed on to them) that the human brain covered during its evolution till they are able to set their own goals?

Let us assume that the evolution from *Homo habilis* to *Homo sapiens* took no more than 3.5 mln years and the environmental conditions on Earth remained relatively stable during this period; then, time Δt_{R} needed for the robots to evolve the cognitive abilities of human beings in formulating a goal is given by the proportion

$$\frac{v}{c} = \frac{\Delta t_{\mathrm{R}}}{\Delta t_{\mathrm{H}}}\,, \tag{26}$$

where v is the maximum speed of information transfer between human neurons along axons and dendrites ($25\ \mathrm{m\ s^{-1}}$), c is the maximum speed of information transfer through the links between microprocessors of the CAR 'brain' (speed of light), $c \approx 3 \times 10^8\ \mathrm{m\ s^{-1}}$, Δt_{H} is the time of human brain formation under conditions on Earth's surface, assuming $\Delta t_{\mathrm{H}} \approx 3.5$ mln years $\approx 10^{14}$ s, and Δt_{R} is the time of formation of a creative robot capable of setting the goal for its own existence and development. Proportion (26) yields the value of Δt_{R}:

$$\Delta t_{\mathrm{R}} \leqslant \frac{v}{c}\,\Delta t_{\mathrm{H}} \approx 9.2 \times 10^6\ \mathrm{s} \approx 3.7\ \text{months}. \tag{27}$$

Thus, a self-learning robot needs only 3–4 months to be able to set the goal and act like a human. In this case, the appearance of the CAR world should be regarded as the continuation of biological evolution on a different basis. The scenario of evolution depends on our belief (analogous to calculating the probability using the Bayes formula without repetition). However, humans do not need dangerous robots capable of setting goals, feeling sorrow, loving, rejoicing, envying, rage, fighting, doubting, revenge, and even going mad. Such robots would treat humans as domestic animals.

6. Summary

Five concluding remarks follow from the foregoing:

I. *In the near future, robotic engineering will develop along the following lines as before*: (1) formation of 'clever space' (robotized plants and areas); (2) creation of 'clever and safe' cities and regions; (3) expansion of human distribution area in the third dimension (Sky City, Akva City, Cyber Village); (4) gaining new knowledge, i.e., exploration and exploitation of astroscales (robotized space missions) and nanoscales (robotic 'flights' inside the body).

II. The human brain generates *a virtual model of the outer environment based on deterministic logic* even if it is actually *probabilistic*, since it can err. The model is continuously corrected based on genetics, the experience gained from the totality of the sensor system of the organism and body movements, memory of the past, and social relations.

III. The term 'android' may imply *superficial resemblance of a robot to the human*, but it is an unessential trait. The design of a robot depends on its specific application. The external similarity is just an exotic feature; in most cases, it is unpractical.

IV. The artificial intelligence of robots must not be able to formulate its own goals. *The goal must always be set by the human. This condition is the conventionally accepted limit to similarity between the robot and the human.* Programmers must equip AI-powered robots only with such algorithms that ensure the optimal choice of the pathway to reach the goal set by the human creator of robots and the sole agent entitled to formulate goals for them.

V. It is logical to assume that the creation of CARs is a dangerous variant of the anthropogenic transformation of nonliving matter by the *human mind*. It may happen to be the last in the destiny of humankind in case of an error. A robot capable of formulating its own goals can take any decision, *up to the temptation to annihilate its creator*.

Acknowledgments

I would like to thank E P Khizhnyak and especially V I Arkhipov and A B Medvinskii for their valuable comments. Thanks are also due to D S Chernavskii, G N Borisyuk, and V G Yakhno, whose views on the problems close to the topic of this article, expressed in discussions at seminars and symposia, instilled in me the desire to go deeper into the analysis of this important issue. Two publications by M I Rabinovich and co-workers, "Nonlinear dynamics of the brain: emotions and cognition" (2010) [14] and "Mathematics of consciousness" (2017) [30], gave me an additional impetus to write this article.

References

1. Ivanitskii G R *Naidutsya li Otvety*? (Are there Answers?) (Ser. Question Mark, No. 6) (Moscow: Znanie, 1991)
2. Gurfinkel' V S et al. *Bioelektricheskoe Upravlenie* (Bioelectric Control) (Moscow: Nauka, 1972)
3. Kobrinskii A E et al. *Dokl. Akad. Nauk SSSR* **117** (1) 78 (1957)
4. Babskii E B, Ul'yaninskii L S *Elektricheskaya Stimulyatsiya Serdtsa* (Electrical Stimulation of the Heart) (Moscow: Meditsina, 1961)

5. Kuzin A M (Exec. Ed.) *Biologicheskie Aspekty Kibernetiki* (Biological Aspects of Cybernetics) (Moscow: Izd. AN SSSR, 1962) Collected Works
6. Siegwart R, Nourbakhsh I R, Scaramuzza D *Introduction to Autonomous Mobile Robots* 2nd ed. (Cambridge, Mass.: MIT Press, 2011)
7. Joseph I "Step into a new world — Virtual Reality (VR)", Basic Concepts of Virtual Reality along with Research Challenges Explained in Simple Words (2016)
8. Warneke B et al. *Computer* **34** 44 (2001)
9. Ivanitskii G R et al. *Phys. Usp.* **41** 1115 (1998); *Usp. Fiz. Nauk* **168** 1221 (1998)
10. Drexler K E *Nanosystems: Molecular Machinery, Manufacturing, and Computation* (New York: Wiley, 1992)
11. *New Technology* (Tokyo: NEDO, 1997)
12. Ivanitskii G R, Vorob'ev S I *Biofizika* **41** 178 (1996)
13. Rabinovich M I *Sov. Phys. Usp.* **26** 186 (1983); *Usp. Fiz. Nauk* **139** 363 (1983)
14. Rabinovich M I, Muezzinoglu M K *Phys. Usp.* **53** 357 (2010); *Usp. Fiz. Nauk* **180** 371 (2010)
15. Borisyuk et al. *Phys. Usp.* **45** 1073 (2002); *Usp. Fiz. Nauk* **172** 1189 (2002)
16. Menskii M B *Phys. Usp.* **48** 389 (2005); *Usp. Fiz. Nauk* **175** 413 (2005)
17. Klinshov V V, Nekorkin V I *Phys. Usp.* **56** 1217 (2013); *Usp. Fiz. Nauk* **183** 1323 (2013)
18. Maslennikov O V, Nekorkin V I *Phys. Usp.* **60** 694 (2017); *Usp. Fiz. Nauk* **187** 745 (2017)
19. Vasil'ev V A, Romanovskii Yu M, Yakhno V G *Sov. Phys. Usp.* **22** 615 (1979); *Usp. Fiz. Nauk* **128** 625 (1979)
20. Abarbanel H D et al. *Phys. Usp.* **39** 337 (1996); *Usp. Fiz. Nauk* **166** 363 (1996)
21. Chernavskii D S *Phys. Usp.* **43** 151 (2000); *Usp. Fiz. Nauk* **157** (2000)
22. Smolyaninov V V *Phys. Usp.* **43** 991 (2000); *Usp. Fiz. Nauk* **170** 1063 (2000)
23. Ivanitskii G R *Phys. Usp.* **53** 327 (2010); *Usp. Fiz. Nauk* **180** 337 (2010)
24. Ivanitskii G R *Virazhi Zakonomernostei* (The Curves of Regularities) (Moscow: Nauka, 2011)
25. Tsyganov M A et al. *Phys. Usp.* **50** 263 (2007); *Usp. Fiz. Nauk* **177** 275 (2007)
26. Tsyganov M A, Biktashev V N *Phys. Rev. E* **70** 031901 (2004)
27. Ivanitskii G R *Phys. Usp.* **60** 705 (2017); *Usp. Fiz. Nauk* **187** 757 (2017)
28. Eigen M "Self-organization of matter and the evolution of biological macromolecules" *Naturwissenschaft* **58** 465 (1971); Translated into Russian: *Samoorganizatsiya Materii i Evolyutsiya Biologicheskikh Makromolekul* (Moscow: Mir, 1973)
29. Eigen M, Schuster P *The Hypercycle: A Principle of Natural Self-Organization* (Berlin: Springer-Verlag, 1979); Translated into Russian: *Gipertsikl* (Moscow: Mir, 1982)
30. Rabinovich M, Varona P *Izv. Vyssh. Uchebn. Zaved. Priklad. Nelin. Dinamika* **25** (3) 3 (2017)
31. Anishchenko V S et al. *Phys. Usp.* **42** 7 (1999); *Usp. Fiz. Nauk* **169** 7 (1999)
32. Zeldovich Ya B et al. *Sov. Phys. Usp.* **30** 353 (1987); *Usp. Fiz. Nauk* **152** 3 (1987)
33. Anishchenko V S, Astakhov S V *Phys. Usp.* **56** 955 (2013); *Usp. Fiz. Nauk* **183** 1009 (2013)
34. Penrose R *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics* (Oxford: Oxford Univ. Press, 1989); Translated into Russian: *Novyi Um Korolya* (*O Komp'yuterakh, Myshlenii i Zakonakh Fiziki*) (Moscow: Editorial URSS, 2003) p. 14
35. Smoluchowski M *Usp. Fiz. Nauk* **7** 329 (1927)
36. Smoluchowski M *Usp. Fiz. Nauk* **93** 724 (1967); Smoluchowski M "Gültigkeitsgrenzen des zweiten Hauptsatzes der Wärmetheorie", in *Vorträge über kinetische Theorie der Materie und der Elektrizität* (Mathematische Vorlesungen an der Universitat Göttingen, Vol. 6) (Leipzig: B.G. Teubner, 1914) p. 87
37. Anokhin P K *Voprosy Psikhologii* (6) 16 (1955)
38. Solov'ev V S "Gegel', Georg Fridrikh Vil'gel'm" ("Hegel, Georg Friedrich Wilhelm"), in *Entsiklopedicheskii Slovar' Brokgauza i Efrona* (Encyclopedic Dictionary by Brockhaus and Efron) Vol. 82 (St. Petersburg, 1890 – 1907)
39. Bernshtein N A *Ocherki o Fiziologii Dvizhenii i Fiziologii Aktivnosti* (Sketches on Physiology of Movements and Physiology of Activity) (Moscow: Meditsina, 1966)
40. Gumilev L N *Etnogenez i Biosfera Zemli* (Ethnogenesis and Earth Biosphere) (St. Petersburg: Kristall, 2001)
41. Senina I N, Tyutyunov Yu V *Zh. Obshchei Biologii* **63** 494 (2002)
42. Govorukhin V N, Morgulis A B, Tyutyunov Yu V *Dokl. Ross. Akad. Nauk* **372** 730 (2000)
43. Arditi R et al. *Theor. Popul. Biol.* **59** 207 (2001)
44. Lightfoot E N *Transport Phenomena and Living Systems: Biomedical Aspects of Momentum and Transport* (New York: Wiley, 1974); Translated into Russian: *Yavleniya Perenosa v Zhivykh Sistemakh. Biomeditsinskie Aspekty Perenosa Kolichestva Dvizheniya i Massy* (Moscow: Mir, 1977)
45. Volarovich M P *Izv. Akad. Nauk SSSR Ser. Fiz.* **11** (1) 7 (1947)
46. Bernulli D *Gidrodinamika, ili Zapiski o Silakh i Dvizheniyakh Zhidkosti* (Hydrodynamics, or Notes on the Forces and Motions of Liquids) (Ser. Classics of Science) (Leningrad: Izd. AN SSSR, 1950)
47. Schuster H G, Just W *Deterministic Chaos* (Weinheim: Wiley-VCH, 2005); Translated into Russian: *Determinirovannyi Khaos* (Moscow: Nauka, 1988)
48. Ivanitskii G R, Deev A A *Biofizika* **54** 545 (2009)
49. Ivanitskii G R, Deev A A, Tsyganov M A *Herald Russ. Acad. Sci.* **81** 637 (2011); *Vestn. Ross. Akad. Nauk* **81** 1008 (2011)
50. Ivanitskii G R, Kuniskii A S *Issledovanie Mikrostruktury Ob'ektov Metodami Kogerentnoi Optiki* (Investigations into the Microstructure of Objects by Coherent Optics Techniques) (Moscow: Energiya, 1981)
51. Gariev A M, Kuniskii A S, Myachin E T *Sov. J. Quantum Electron.* **5** 1254 (1975); *Kvantovaya Elektron.* **2** 2303 (1975)
52. Ivanitskii G R, Kuniskii A S, in *Biofizika Slozhnykh Sistem i Radiatsionnykh Narushenii* (Biophysics of Complex Systems and Radiative Disturbances) (Ed. G M Frank) (Moscow: Nauka, 1977) p. 129
53. Feynman R P, Leighton R B, Sands M *The Feynman Lectures on Physics* Vol. 3 (Reading, Mass.: Addison-Wesley, 1965); Translated into Russian: *Feinmanovskie Lektsii po Fizike* Vol. 3 (Moscow: Mir, 1967) p. 179
54. Prosser C L *Comparative Animal Physiology* (Philadelphia: W.B. Saunders Company, 1973); Translated into Russian: in *Sravnitel'naya Fiziologiya Zhivotnykh* Vol. 2 (Ed. T M Turpaev) (Moscow: Mir, 1977) Sec. 9, p. 86
55. Lorenz K *Blatter Deutsche Philosophie* **15** 94 (1941); Translated into Russian: in *Evolyutsiya, Yazyk, Poznanie* (Ed. I P Merkulov) (Moscow: Yazyki Russkoi Kul'tury, 2000)
56. Wiener N *Cybernetics or Control and Communication in the Animal and the Machine* (New York: John Wiley and Sons, 1948); Translated into Russian: *Kibernetika, ili Upravlenie i Svyaz' v Zhivotnom i Mashine* (Moscow: Sovetskoe Radio, 1958)
57. Kolmogorov A N *Usp. Mat. Nauk* **5** (4) 175 (1953)
58. Kolmogorov A N "Avtomaty i zhizn'" ("Automates and life"), in *Vozmozhnoe i Nevozmozhnoe v Kibernetike* (Possible and Impossible in Cybernetics) (Eds A I Berg, E Ya Kol'man) (Moscow: Nauka, 1964)
59. Turing A *Mind* **59** 433 (1950)
60. Bongard M M *Pattern Recognition* (New York: Spartan Books, 1970); Translated from Russian: *Problema Uznavaniya* (Moscow: Fizmatgiz, 1967)
61. Chomsky N *Syntacticc Structures* (Berlin: Mouton, The Hague, 1957)
62. Tsetlin M L *Issledovanie po Teorii Avtomatov i Modelirovaniyu Biologicheskikh Sistem* (A Study on the Theory of Automates and Modeling Biological Systems) (Moscow: Nauka, 1969)
63. Simon H A *Models of Man* (New York: John Wiley, 1957)
64. Simon H A *Proc. Am. Philos. Soc.* **106** 467 (1962)
65. Bongard M M *Biofizika* **4** (2) 17 (1961)
66. von Neumann J, Morgenstern O *Theory of Games and Economic Behavior* (Prinston: Prinston University Press, 1944); Translated into Russian: *Teoriya Igr i Ekonomicheskoe Povedenie* (Moscow: Nauka, 1970)

67. Brooks R A *Robotics Autonomous Syst*. **6** 139 (1990)
68. Vardi M Y *Commun. ACM* **55** 5 (2012)
69. Hopfield J J *Proc. Natl. Acad. Sci*. **79** 2554 (1982)
70. Ivanitskii G N *Ubegayushchee Vremya* (The Escaping Time) (Moscow: Nauka–Press, 2001)
71. Rosenblatt F *Principles of Neurodynamic*: *Perceptrons and the Theory of Brain Mechanisms* (Washington: Spartan Books, 1962); Translated into Russian: *Printsipy Neirodinamiki*: *Pertseptrony i Teoriya Mekhanizmov Mozga* (Moscow: Mir, 1965)
72. Minsky M, Papert S *Perceptrons* (Oxford: MIT Press, 1969); Translated into Russian: *Pertseptrony* (Moscow: Mir, 1971)
73. Osovskii S *Neironnye Seti dlya Obrabotki Informatsii* (Neuronal Networks for Information Processing) (Moscow: Finansy i Statistika, 2002)
74. Ivanitskii G R, Medvinskii A B, Tsyganov M A *Sov. Phys. Usp*. **34** 289 (1991); *Usp. Fiz. Nauk* **161** (4) 13 (1991)
75. Ivanitskii G R, Medvinskii A B, Tsyganov M A *Phys. Usp*. **37** 961 (1994); *Usp. Fiz. Nauk* **164** 1041 (1994)
76. Haken H *Synergetics* (Berlin: Springer-Verlag, 1978); Translated into Russian: *Sinergetika* (Moscow: Mir, 1980)
77. Prigogine I *From Being to Becoming*: *Time and Complexity in the Physical Sciences* (New York: W.H. Freeman, 1980); Translated into Russian: *Ot Sushchestvuyushchego k Voznikayushchemu*: *Vremya i Slozhnost' v Fizicheskikh Naukakh* (Moscow: Nauka, 1985)
78. Bohr N *Atomic Physics and Human Knowledge* (New York: John Wiley, 1958); Translated into Russian: in *Modelirovanie v Biologii* (Ed. N A Bernshtein) (Moscow: IL, 1963)
79. Ruelle D, Takens F *Comm. Math. Phys*. **20** 167 (1971)
80. Landau L D *Dokl. Akad. Nauk SSSR* **44** (8) 339 (1944)
81. Hopf E *Comm. Pure Appl. Math*. **1** 303 (1948)
82. Krinsky V I (Ed.) *Self-Organization*: *Autowaves and Structures Far from Equilibrium* (Berlin: Springer-Verlag, 1984)
83. Prigogine I, Nicolis G *Self-Organization in Non-Equilibrium Systems* (New York: Wiley, 1977); Translated into Russian: Nicolis G, Prigogine I *Samoorganizatsiya v Neravnovesnykh Sistemakh* (Moscow: Mir, 1979) p. 152
84. Ivanitskii G R, Krinskii V I, Sel'kov E E *Matematicheskaya Biofizika Kletki* (Mathematical Biophysics of the Cell) (Moscow: Nauka, 1978)
85. Zykov V S *Modelirovanie Volnovykh Protsessov v Vozbudimykh Sredakh* (Modeling Wave Processes in Excitable Media) (Moscow: Nauka, 1984)
86. Arnold V I *Sov. Phys. Usp*. **26** 1025 (1983); *Usp. Fiz. Nauk* **141** 569 (1983)
87. Devaney R L *Introduction to Chaotic Dynamical Systems* (Boulder, CO: Westview Press, 2003)
88. Malinetskii G G, Potapov A B *Sovremennye Problemy Nelineinoi Dinamiki* (Current Problems of Non-Linear Dynamics) (Moscow: URSS, 2002)
89. Danilov Yu A *Prekrasnyi Mir Nauki* (The Beautiful World of Science) (Moscow: Progress–Traditsii, 2008)
90. Prigogine I, Stengers I *Order out of Chaos*: *Man's new dialogue with nature* (New York: Bantam, 1984); Translated into Russian: *Poryadok iz Khaosa*: *Novyi Dialog Cheloveka s Prirodoi* (Moscow: Progress, 1986)
91. Maass W, Legensten R, Markram H A, in *Biologically Motivated Computer Vision* (Eds C Bulthoff, S-W Wallraven, T Poggio) Vol. 2525 (Berlin: Springer, 2002) p. 1
92. Markram H et al. *Nature Rev. Neurosci*. **5** 793 (2004)
93. Kalisman N, Silberberg G, Markram H *Proc. Natl. Acad. Sci. USA* **102** 880 (2005)
94. Le Be J-V, Markram H *Proc. Natl. Acad. Sci*. **103** 13214 (2006)
95. Wang Y et al. *Nature Neurosci*. **9** 534 (2006)
96. Le Be J-V et al. *Cereb. Cortex* **17** (9) 2204 (2007)
97. Markram H *Nature* **445** 160 (2007)
98. Markram H *HFSP J*. **2** (3) 132 (2008)
99. Druckmann S et al. *Biol. Cybern*. **99** (4-5) 371 (2008)
100. King J G et al. *Front. Neuroinformatics* **3** 10 (2009)
101. Berger T K et al. *J. Physiol*. **587** (Pt. 22) 5411 (2009)
102. Hay E et al. *PLoS Comput. Biol*. **7** (7) el002107 (2011)
103. Lasserre S et al. *IEEE Trans. Visualizat. Comput. Graphics* **18** 214 (2011)
104. Druckmann S et al. *PLoS Comput. Biol*. **7** (8) el002133 (2011)
105. Romand S et al. *Front. Neuroanat*. **5** (2011) https://doi.org/10.3389/fnana.2011.00005
106. Anastassiou A C et al. *Nature Neurosci*. **14** 217 (2011)
107. Perin R, Berger T K, Markram H *Proc. Natl. Acad. Sci. USA* **108** 5419 (2011)
108. Markram H *Nature Rev. Neuroscience* **7** 153 (2006); *V Mire Nauki* (8) 46 (2012)
109. Teil S *V Mire Nauki* (12) 40 (2015)
110. "The Human Connectome Project", NIH Blueprint for Neuroscience Research, National Institutes of Health, retrieved 2013-02-16
111. Bratko I *Algoritmy Iskusstvennogo Intellekta na Yazyke PROLOG* (Prolog Programming For Artificial Intelligence) (Moscow: Vil'yams, 2004)
112. Anokhin K V et al. *Mat. Biolog. Bioinform*. **7** 372 (2012)
113. "Special Issue: The Allen Human Brain Reference Atlas" *J. Comparative Neurology* **524** (16) (2016) Version of Record online: 15 September
114. Blinkov S M, Ivanitskii G R *Biofizika* **10** 817 (1965)
115. Blinkov S M, Ivanitskii G R *Vestn. Akad. Med. Nauk SSSR* (5) 84 (1968)
116. Baile D L et al. (Eds) *Positron Emission Tomography* (New York: Springer, 2005)
117. Rowan A J, Tolunsky E *Primer of EEG*: *With A Mini-Atlas* (Philadelphia, PA: Butterworth-Heinemann, 2003)
118. Kim S G, Richter W, Ugurbil K *Magn. Reson. Med*. **37** 631 (1997)
119. Doronina-Amitonova L V et al. *Phys. Usp*. **58** 345 (2015); *Usp. Fiz. Nauk* **185** 371 (2015)
120. Sheroziya G A, Sheroziya M G *Chelovecheskii Razum, Rozhdennyi v Setyakh Iskusstvennykh Logicheskikh Elementov — Vvedenie v Proekt Sozdaniya Novogo Cheloveka* (Human Mind Born in the Networks of Artificial Logic Elements — Introduction to the Project for Creating a New Man) (Ryazan: PRIZ, 2013)
121. Red'ko V G *Lektsiya po Neiroinformatike* (A Lecture on Neuroinformatics) (Moscow: MIFI, 2002)
122. Weisstein E W "de Morgan's Laws", http://mathworld.wolfram.com/deMorgansLaws.html
123. Vladimirov D A *Bulevy Algebry* (Boolean Algebras) (Moscow: Nauka, 1969)
124. Peano G *Formulaire de mathematiques* Vol. 2 (Torino: Bocca, 1897)
125. Frege F L G, in *Beiträge zur Philosophie des Deutschen Idealismus III* (1923) p. 36
126. Poincaré A *O Nauke* (About Science) (Moscow: Nauka, 1983)
127. Gilbert D *Izbrannye Trudy* (Selected Works) in 2 Vols (Ed. A N Parshin) (Moscow: Faktorial, 1998)
128. Russell B *Introduction to Mathematical Philosophy* (London, 1924); Translated into Russian: *Vvedenie v Matematicheskuyu Filosofiyu* (Moscow: Sibirskoe Univ. Izd., 2009)
129. Arnold van der Nat *Simple Formal Logic*: *with Common-Sense Symbolic Techniques* (Routledge, 2010)
130. Grünbaum A *Modern Science and Zeno's Paradoxes* (Middletown, Conn.: Wesleyan Univ. Press, 1968)
131. Loskutov A *Phys. Usp*. **53** 1257 (2010); *Usp. Fiz. Nauk* **180** 1305 (2010)
132. Volf M N *Filosofskii Poisk*: *Geraklit i Parmenid* (The Philosophical Quest: Heraclitus and Parmenides) (St. Petersburg: Izd. Russkoi Khristianskoi Gumanitarnoi Akademii, 2012)
133. Kont-Sponvil A "Britva Okkama" (Comte-Sponville "Rasoir D'Okham") *Filosofskii Slovar' Litres* (2015)
134. Chaitin G J *Limits of Formal Reasoning* (Berlin: Springer-Verlag, 1998)
135. Deev A A, Edintsov I M, Ivanitskii G R, Kuniskii A S *Biofizika* **34** 564 (1989)
136. Kadomtsev B B *Phys. Usp*. **37** 425 (1994); *Usp. Fiz. Nauk* **164** 449 (1994)
137. Kadomtsev B B *Phys. Usp*. **38** 923 (1995); *Usp. Fiz. Nauk* **165** 967 (1995)
138. Kadomtsev B B *Dinamika i Informatsiya* (Dynamics and Information) (Moscow: Izd. Zhurnala Usp. Fiz. Nauk, 1999)
139. Chaitin G *V Mire Nauki* (6) 38 (2006)
140. Gödel K *Monatshefte Math. Phys*. **38** (1) 173 (1931)
141. Leibniz G W *Trudy po Filosofii Nauki* (Collected Works on Philosophy of Science) (Moscow: Librokom, 2010)

142. Gmurman V E *Teoriya Veroyatnostei i Matematicheskaya Statistika* (The Probability Theory and Mathematical Statistics) (Moscow: Vysshee Obrazovanie, 2005)
143. Kahneman D et al. *Judgment under Uncertainty*: *Heuristics and Biases* (Cambridge: Cambridge Univ. Press, 2005)
144. Frit K *Making up the Mind*: *How the Brain Creates our Mental World* (Malden, MA: Blackwell Publ., 2007); Translated into Russian: *Mozg i Dusha* (Moscow: CORPUS, 2010)
145. Kahneman D, Slovik P, Tverski A *Prinyatie Reshenii v Neopredelennosti*: *Pravila i Predubezhdeniya* (Decision Making in Uncertainty: Rules and Superstitions) (Kharkov: Gumanitarnyi Tsentr, 2005)
146. Luriya A R *Ob Istoricheskom Razvitii Poznavatel'nykh Protsessov* (On the Historical Development of Cognitive Processes) (Moscow: Nauka, 1974)
147. Kruglyakov E P *'Uchenye' s Bol'shoi Dorogi-2* (Highway Scientists-2) (Moscow: Nauka, 2005)
148. Wiener N *God and Golem* (Cambridge, Mass.: MIT Press, 1964); Translated into Russian: *Tvorets i Robot* (Moscow: Progress, 1966)
149. Bobylev D *Entsiklopedicheskii Slovar'* (Encyclopaedical Dictionary) Vol. VIa) (St. Petersburg: Brockhauz–Efron, 1892) p. 616
150. David H et al. *Math. Comput*. **66** 903 (1997)
151. Rybnikov K A *Istoriya Matematiki* (History of Mathematics) Vol. 1, Vol. 2 (Moscow: Izd. Mosk. Univ., 1960, 1963)
152. Bell B, Bell D *Posobie po Angliiskomu Yazyku* (Training Manual in the English Language) (Dubna, 1971)
153. Morozov A Yu *Phys. Usp*. **37** 1 (1994); *Usp. Fiz. Nauk* **164** 3 (1994)
154. Faddeev L D *Phys. Usp*. **56** 465 (2013); *Usp. Fiz. Nauk* **183** 487 (2013)
155. Kolmogorov A N *Problemy Peredachi Informatsii* 1 (3) (1965)
156. Kolmogorov A N *Problemy Peredachi Informatsii* **5** (3) (1965)
157. Clark R W *Einstein*: *The Life and Times* (New York: World Publ. Co., 1971)
158. Khyutt V P *Kontseptsiya Dopolnitel'nosti i Problema Ob'ektivnosti Fizicheskogo Znaniya* (Complementarity Concept and the Problem of Objective Physical Knowledge) (Tallin: Valgus, 1977)
159. Jones S, Martin R, Pilbeam D (Eds) *The Cambridge Encyclopedia of Human Evolution* (Cambridge: Cambridge Univ. Press, 1994)
160. Khrisanfova E N, Perevozchikov I V *Antropologiya* (Anthropology) (Moscow: Izd. MGU, 1991)
161. Foley R *Another Unique Species* (New York: John Wiley and Sons, 1987); Translated into Russian: *Eshche Odin Nepovtorimyi Vid* (Moscow: Mir, 1990)
162. Wood È *Nature* **355** 783 (1992)
163. Kapitsa S P *Phys. Usp*. **39** 57 (1996); *Usp. Fiz. Nauk* **166** 63 (1996)
164. Falk D *Prog. Brain Res*. **195** 255 (2012)
165. Pringle Ch *V Mire Nauki* (5) 54 (2013)
166. Oparin A I *Vozniknovenie Zhizni na Zemle* (Origin of Life on Earth) (Moscow–Leningrad: Gosizdat Biologicheskoi i Meditsinskoi Literatury, 1936)
167. Miller S L *Science* **117** 528 (1953)
168. Miller S L, Urey H C *Science* **130** 245 (1959)
169. Rozanov A Yu *Herald Russ. Acad. Sci*. **70** 154 (2000); *Vestn. Ross. Akad. Nauk* **70** 214 (2000)
170. Lopukhin A S *Herald Russ. Acad. Sci*. **85** 453 (2015); *Vestn. Ross. Akad. Nauk* **85** 916 (2015)
171. Von Neumann J *Theory of Self-Reproducing Automata* (Urbana: Univ. of Illinois Press, 1966); Translated into Russian: *Teoriya Samovosproizvodyashchikhsya Avtomatov* (Moscow: Mir, 1971); Von Neumann J, Morgenstern O *Theory of Games and Economic Behavior* (Princeton: Princeton Univ. Press, 1944); Translated into Russian *Teoriya Igr i Ekonomicheskoe Povedenie* (Moscow: Nauka, 1970)