Available online at www.sciencedirect.com

SCIENCE DIRECT®

Physics of Life Reviews 3 (2006) 23–55

PHYSICS of LIFE reviews

www.elsevier.com/locate/plrev

Review

Toward physics of the mind: Concepts, emotions, consciousness, and symbols

Leonid I. Perlovsky *

Air Force Research Lab., 80 Scott Rd., Hanscom AFB, MA 01731, USA

Accepted 28 November 2005

Available online 20 January 2006

Abstract

Mathematical approaches to modeling the mind since the 1950s are reviewed, including artificial intelligence, pattern recognition, and neural networks. I analyze difficulties faced by these algorithms and neural networks and relate them to the fundamental inconsistency of logic discovered by Gödel. Mathematical discussions are related to those in neurobiology, psychology, cognitive science, and philosophy. Higher cognitive functions are reviewed including concepts, emotions, instincts, understanding, imagination, intuition, consciousness. Then, I describe a mathematical formulation, unifying the mind mechanisms in a psychologically and neuro-biologically plausible system. A mechanism of the knowledge instinct drives our understanding of the world and serves as a foundation for higher cognitive functions. This mechanism relates aesthetic emotions and perception of beauty to "everyday" functioning of the mind. The article reviews mechanisms of human symbolic ability. I touch on future directions: joint evolution of the mind, language, consciousness, and cultures; mechanisms of differentiation and synthesis; a manifold of aesthetic emotions in music and differentiated instinct for knowledge. I concentrate on elucidating the first principles; review aspects of the theory that have been proven in laboratory research, relationships between the mind and brain; discuss unsolved problems, and outline a number of theoretical predictions, which will have to be tested in future mathematical simulations and neuro-biological research.

Keywords: The mind; Physics; Emotions; Concepts; Consciousness; The knowledge instinct

Contents

1. Physics and the mind 24
2. Logic vs. mind 25
3. Computational intelligence since the 1950s: Complexity and logic 26
4. Structures of the mind 27
5. Modeling field theory (MFT) 30
   5.1. The knowledge instinct 30
   5.2. Dynamic logic 32
   5.3. Example of dynamic logic operations 33
   5.4. MFT hierarchical organization 33

* Tel. +781 377 1728.
E-mail address: leonid.perlovsky@hanscom.af.mil.

1571-0645/$ – see front matter
doi:10.1016/j.plrev.2005.11.003

6. Symbolic culture . . . 35
6.1. Symbols in cognitive linguistics, philosophy, and computational intelligence . . . 35
6.2. Integration of language and cognition . . . 38
7. Workings of the mind: Current understanding, future research, predictions and tests . . . 40
7.1. Cognition and dynamics of MFT . . . 40
7.2. Elementary thought-process . . . 40
7.3. Understanding . . . 41
7.4. Conscious and unconscious . . . 41
7.5. Imagination . . . 45
7.6. Bodily instincts and emotions . . . 46
7.7. Aesthetic emotions and the instinct for knowledge . . . 46
7.8. Beautiful and sublime . . . 47
7.9. Intuition . . . 47
7.10. Language, cognition, and symbols . . . 48
7.11. Creativity, differentiation, and synthesis . . . 49
7.12. Teleology, causality, and the knowledge instinct . . . 50
7.13. Mind and brain, experimental evidence . . . 50
7.14. Predictions and testing . . . 50
Acknowledgements . . . 51
References . . . 51

1. Physics and the mind

Is a physical theory of the mind possible? What kind of physics would this be? I proceed assuming that the mind and brain refer to the same physical system at different levels of description. This situation is not new to physics, for example thermodynamics and statistical physics are also related to each other in this manner. Although a physicist would prefer the most fundamental description level (elementary particles, strings ...), intermediate levels are sometimes appropriate. Einstein once mentioned that he liked thermodynamics as physics defined at an intermediate phenomenological level. The world is amenable to understanding at various levels. Understanding searched by physicists is specific in certain ways: physics is a search for basic laws, a few universal "first principles" describing a wealth of observed phenomena.

Many physicists are uncomfortable with the phrase "physics of the mind", and I will attempt to overcome this initial reaction. Some of the reasons for discomfort are obvious: the mind is perceived as deeply personal, something that no equation will ever be able to describe, no computer will ever be able to simulate. Responding to this reservation, let me mention that no particular individual mind is addressed here, rather this review considers the most general mechanisms that act in every mind. The future will tell how close a physical theory could come to understanding individual minds. Another reason for skepticism is that the mind is both diverse and unpredictable, therefore how can it be reduced to few basic laws? Newton saw nothing wrong with developing physics of the mind, which he called spiritual substance. However Newton failed and since then few physicists have dared to approach the subject. Recently, new data, new intuitions, and new mathematical tools have emerged, and today we make a new attempt. We seek to identify a few basic principles of the mind operation, formulate these principles mathematically, use them to explain a wealth of known data, and make predictions that can be tested in the lab.

How the mind works has been the subject of discussions for millennia, from the Ancient Greek philosophers to mathematicians and cognitive scientists of today. Words like *mind, thought, imagination, emotion, concept* present a challenge: people use these words in many ways colloquially, but in cognitive science and in mathematics of intelligence they have not been uniquely defined and their meaning is a subject of active research and ongoing debates [1]. Standardized definitions come at the end of the development of a theory (e.g., "force" was defined by Newton's laws, following centuries of less precise usage). Whereas the mind theory is under development, this review adheres to the following guidance: we need to make sure that our proposals:

(1) correspond to the discussions in scientific and mathematical community,
(2) correspond to millennia of philosophical discussions and the general cultural usage,

(3) are clear and mathematically tractable, and
(4) that deviations or discrepancies in these proposals are noted and discussed.

According to the dictionary [2], which we take as a starting point, the mind includes conscious and unconscious processes, especially thought, perception, emotion, will, memory, and imagination, and it originates in brain. These constituent notions will be discussed throughout the paper. Specific neural mechanisms in the brain "implementing" various mind functions constitute the relationship between the mind and brain; this is a contemporary formulation of the millennia old mind-body problem. We will discuss possible relationships between the proposed mathematical descriptions and neural structures in the brain.

A broad range of opinions exists about the mathematical methods suitable for the description of the mind. Founders of artificial intelligence, including Allan Newell and Marvin Minsky, thought that formal logic was sufficient [3] and that no specific mathematical techniques would be needed to describe the mind [4]. An opposite view was advocated by Brian Josephson and Roger Penrose, suggesting that the mind cannot be understood within the current knowledge of physics; new unknown yet physical phenomena will have to be accounted for explaining the working of the mind [5,6]. Quantum computational processes were considered, which might take place in the brain [5,7,8]. Some authors are developing a "classical physics" point of view in which there are few specific mathematical constructs, "the first principles" of the mind. Among researchers taking this view are Grossberg, who suggests that the first principles include a resonant matching between bottom-up signals and top-down representations [9], and emotional evaluation of conceptual contents [10]; Zadeh developing theory of granularity [11]; Meystel developing hierarchical multi-scale organization [12]; Edelman suggesting neuronal group selection [13]; and the author, suggesting the knowledge instinct, aesthetic emotions, and dynamic logic [8,14,15] among the first principles of the mind. This review addresses mathematical methods that can be realized by classical physics mechanisms. We review specific difficulties encountered by previous attempts at mathematical modeling of the mind and recent developments overcoming these difficulties.

2. Logic vs. mind

For a long time people believed that intelligence is equivalent to conceptual understanding and reasoning. A part of this belief was that the mind works according to logic. Although it is obvious that the mind is not logical, over the course of the two millennia since Aristotle, many people have identified the power of intelligence with logic. Founders of artificial intelligence in the 1950s and 1960s believed that by relying on rules of logic they would soon develop computers with intelligence far exceeding the human mind.

The beginning of this story is usually attributed to Aristotle, the inventor of logic [16]. However, Aristotle did not think that the mind works logically; he invented logic as a supreme way of argument, not as a theory of the mind. This is clear from many Aristotelian writings, for example, in "Rhetoric for Alexander" Aristotle lists dozens of topics on which Alexander had to speak publicly [17]. For each topic, Aristotle identified two opposite positions (e.g. make peace or declare war; use torture or do not for extracting the truth, etc.). For each of the opposite positions, Aristotle gives logical arguments, to argue either way. Clearly, for Aristotle, logic is a tool to express previously made decisions, not the mechanism of the mind. Logic can only provide deductions from first principles, but cannot indicate what the first principles should be. Logic, if you wish, is a tool for politicians. (Scientists, I would add, use logic to present their results, but not to arrive at these results.) To explain the mind, Aristotle developed a theory of Forms, which will be discussed later. But during the following centuries the subtleties of Aristotelian thoughts were not always understood. With the advent of science, the idea that intelligence is equivalent to logic was gaining grounds. In the 19th century mathematicians turned their attention to logic. George Boole noted what he thought was not completed in Aristotle's theory. The foundation of logic, since Aristotle, was a law of excluded middle (or excluded third): every statement is either true or false, any middle alternative is excluded [18]. But Aristotle also emphasized that logical statements should not be formulated too precisely (say, a measure of wheat should not be defined with an accuracy of a single grain), that language implies the adequate accuracy, and everyone has his mind to decide what is reasonable.

Boole thought that the contradiction between exactness of the law of excluded middle and vagueness of language should be corrected. A new branch of mathematics, *formal logic* was born. Prominent mathematicians contributed to the development of formal logic, including George Boole, Gottlob Frege, Georg Cantor, Bertrand Russell, David Hilbert, and Kurt Gödel. Logicians "threw away" uncertainty of language and founded formal mathematical logic based on the law of excluded middle. Most of physicists today agree that exactness of mathematics is an inseparable part of physics, but formal logicians went beyond this. Hilbert developed an approach named formalism, which rejected the intuition as a part of scientific investigation and thought to define scientific objects formally in terms of axioms or rules. Hilbert was sure that his logical theory also described mechanisms of the mind: "The fundamental idea of my proof theory is none other than to describe the activity of our understanding, to make a protocol of the rules according to which our thinking actually proceeds" [19]. In the 1900 he formulated famous Entscheidungsproblem: to define a set of logical rules sufficient to prove all past and future mathematical theorems. This entailed formalization of scientific creativity and the entire human thinking.

Almost as soon as Hilbert formulated his formalization program, the first hole appeared. In 1902 Russell exposed an inconsistency of formal procedures by introducing a set R as follows: *R is a set of all sets which are not members of themselves*. Is R a member of R? If it is not, then it should belong to R according to the definition, but if R is a member of R, this contradicts the definition. Thus, either way we get a contradiction. This became known as the Russell's paradox. Its joking formulation is as follows: A barber shaves everybody who does not shave himself. Does the barber shave himself? Either answer to this question (yes or no) leads to a contradiction. This barber, like Russell's set can be logically defined, but cannot exist. For the next 25 years mathematicians where trying to develop a self-consistent mathematical logic, free from the paradoxes of this type. But, in 1931, Gödel has proved that it is not possible [20], formal logic was inconsistent, self-contradictory.

Belief in logic has deep psychological roots related to functioning of human mind. A major part of any perception and cognition process is not accessible to consciousness directly. We are conscious about the "final states" of these processes, which are perceived by our minds as "concepts" approximately obeying formal logic. For this reason prominent mathematicians believed in logic. Even after the Gödelian proof, founders of artificial intelligence still insisted that logic is sufficient to explain working of the mind. We will turn to this in the next section; for now, let us just state that logic is not a fundamental mechanism of the mind, but the result of mind's operations (in Section 5 we discuss mathematics of dynamic logic, which suggests a mathematical explanation of how logic appears from illogical states).

3. Computational intelligence since the 1950s: Complexity and logic

Simple object perception involves signals from sensory organs and internal mind's representations (memories) of objects. During perception, the mind associates subsets of signals corresponding to objects with representations of object. This produces object recognition; it activates brain signals leading to mental and behavioral responses, parts of understanding.

Developing mathematical descriptions of the very first *recognition* step in this seemingly simple association–recognition–understanding process has not been easy, a number of difficulties have been encountered during the past fifty years. These difficulties were summarized under the notion of combinatorial complexity (CC) [21]. CC refers to multiple combinations of various elements in a complex system; for example, recognition of a scene often requires concurrent recognition of its multiple elements that could be encountered in various combinations. CC is prohibitive because the number of combinations is very large: for example, consider 100 elements (not too large a number); the number of combinations of 100 elements is 100^{100}, exceeding the number of all elementary particle events in life of the Universe; no computer would ever be able to compute that many combinations.

The problem was first identified in pattern recognition and classification research in the 1960s and was named "the curse of dimensionality" [22]. It seemed that adaptive self-learning algorithms and neural networks could learn solutions to any problem "on their own", if provided with a sufficient number of training examples. The following thirty years of developing adaptive statistical pattern recognition and neural network algorithms led to a conclusion that the required number of training examples often was combinatorially large. Thus, self-learning approaches encountered *CC of learning requirements*.

Rule-based systems were proposed in the 1970s to solve the problem of learning complexity [23,24]. An initial idea was that rules would capture the required knowledge and eliminate a need for learning. However in presence of variability, the number of rules grew; rules became contingent on other rules; combinations of rules had to be considered; rule systems encountered *CC of rules*.

Beginning in the 1980s, model-based systems were proposed. They used models which depended on adaptive parameters. The idea was to combine advantages of rules with learning–adaptivity by using adaptive models. The knowledge was encapsulated in models, whereas unknown aspects of particular situations were to be learned by fitting model parameters [25,26]. Fitting models to data required selecting data subsets corresponding to various models. The number of subsets, however, is combinatorially large. A general popular algorithm for fitting models to the data, multiple hypothesis testing [27], is known to face CC of computations. Model-based approaches encountered *computational CC* (N and NP complete algorithms).

In subsequent research, CC was related to the type of logic, underlying various algorithms and neural networks [21]. Formal logic is based on the "law of excluded middle", according to which every statement is either true or false and nothing in between. Therefore, algorithms based on formal logic have to evaluate every little variation in data or internal representations as a separate logical statement (hypothesis); a large number of combinations of these variations causes combinatorial complexity. In fact, combinatorial complexity of algorithms based on logic is related to Gödel theory: it is a manifestation of the inconsistency of logic in finite systems [28]. Multivalued logic and fuzzy logic were proposed to overcome limitations related to the law of excluded third [29]. Yet the mathematics of multivalued logic is no different in principle from formal logic, "excluded third" is substituted by "excluded $n + 1$". Fuzzy logic encountered a difficulty related to the degree of fuzziness, if too much fuzziness is specified, the solution does not achieve a needed accuracy, if too little, it will become similar to formal logic. Complex systems require different degrees of fuzziness in various elements of system operations; searching for the appropriate degrees of fuzziness among combinations of elements again would lead to CC. Is logic still possible after Gödel? Bruno Marchal recently reviewed the contemporary state of this field [30], it appears that logic after Gödel is much more complicated and much less logical than was assumed by founders of artificial intelligence. Also, CC is still unsolved within logic. Penrose thought that Gödel's results entail incomputability of the mind processes and testify for a need for new physics [31]. An opposite position in this review is that incomputability of logic does not entail incomputability of the mind. Logic is not the basic mechanism of the mind.

Various manifestations of CC are all related to formal logic and Gödel theory. Rule systems rely on formal logic in a most direct way. Self-learning algorithms and neural networks rely on logic in their training or learning procedures: every training example is treated as a separate logical statement. Fuzzy logic systems rely on logic for setting degrees of fuzziness. CC of mathematical approaches to the mind is related to the fundamental inconsistency of logic.

4. Structures of the mind

In the 1950s and 1960s developers of artificial intelligence naïvely believed that they would soon develop computers exceeding human intelligence, and that the mathematics of logic was sufficient for this purpose. As we discussed, logic does not work, but the mind works. So let us turn to the mechanisms of the mind discussed in psychology, philosophy, cognitive science, and neurobiology. Possibly, we will find inspiration for developing mathematics needed for physics of the mind and for intelligent computers. The main mechanisms of the mind include instincts, concepts, emotions, and behavior. Let us look briefly at their current definitions in cognitive science and psychology.

While theories of life and intelligence are being developed, as mentioned, definitions of cognitive functions are a subject of research and debates. Let me summarize few related definitions [2,32,33] as a starting point for further elaboration. Instincts are innate capabilities, aptitudes, or behavior, which are complex, not learned, and normally adaptive. In humans and higher animals, instincts are related to emotions.

We use the word "concept" to designate a common thread among words like concept, idea, understanding, thought, or notion. Different authors use these words with different meanings or subtle differences. A common thread among these words is an abstract, universal psychical entity that serves to designate a category or class of entities, events, or relations. Concepts are abstract in that they treat individual entities as if they were identical. Emphasizing this property, Middle Age philosophers used the term "universals". Plato and Aristotle called them ideas or forms, and considered them the basis for the mind's understanding of the world. Similarly, Kant considered them a foundation for the ability for understanding, the contents of pure reason. According to Jung, conscious concepts of the mind are learned on the basis of inborn unconscious psychic structures, archetypes. Contemporary science often equates the mechanism of concepts with internal representations of objects, their relationships, situations, etc.

Ray Jackendoff [34] considers the terms *representation* or *symbol* as too loaded with "thorny philosophical problem of intentionality", and uses the word *model*. I do not think we should be afraid of intentionality; John Searle emphasis on intentionality as "aboutness" [35,36] is too narrow [37]. All brain mechanisms and mental functions are intentional; in fact everything within a living being is a result of long evolution and has evolved with a certain intent, or better to say, purpose. We are purposeful beings, and I will return to this discussion later. But I agree with Jackendoff that the word model is most appropriate for concept or representation.

Emotions refer to both exaggeratedly expressive communications and to internal states related to feelings. Love, hate, courage, fear, joy, sadness, pleasure and disgust can all be described in both psychological and physiological terms. Emotion is the realm where thought and physiology are inextricably entwined, and where the self is inseparable from individual perceptions of value and judgment toward others and ourselves. Emotions are sometimes regarded as the antithesis of reason; as is suggested by phrases such as "appeal to emotion" or "don't let your emotions take over". A distinctive and challenging fact about human beings is a potential for both opposition and entanglement between will, emotion, and reason. It has also been suggested that there is no empirical support for any generalization suggesting the antithesis between reason and emotion: indeed, anger or fear can often be thought of as a systematic response to observed facts. What should be noted, however, is that the human psyche possesses many possible reactions and perspective in regard to the internal and external world—often lying on a continuum—some of which may involve the extreme of pure intellectual logic (often called "cold"), other the extreme of pure emotion unresponsive to logical argument ("the heat of passion"). In any case, it should be clear that the relation between logic and argument on the one hand and emotion on the other, is still a matter of research. It has been noted by many that passion, emotion, or feeling can add backing to an argument, even one based primarily on reason—particularly in regard to religion or ideology, areas of human thought which frequently demand an all-or-nothing rejection or acceptance, that is, the adoption of a comprehensive worldview partly backed by empirical argument and partly by feeling and passion. Moreover, it has been suggested by several researchers (see e.g., D.S. Levine and S.J. Leven [38]) that typically there is no "pure" decision or thought, that is, no thought based "purely" on intellectual logic or "purely" on emotion—most decisions and cognitions are founded on a mixture of both.

An essential role of emotions in working of the mind was analyzed by many researchers, from various perspectives: philosophical (Rene Descartes [39], Immanuel Kant [40], Jean Paul Sartre [41]); analytical psychology (Carl Jung [42]); psychological and neural (Stephen Grossberg and Daniel Levine [10], Andrew Ortony [43], Joseph LeDoux [44]); philosophical–linguistic (Paul Griffiths [45]); neuro-physiological (Antonio Damasio [46]); and from the learning and cognition perspective by the author [47]. Descartes attempted a scientific explanation of passions. He rationalized emotions, explained them as objects, and related to physiological processes. According to Kant, emotions are closely related to judgments about which individual experiences and perceptions correspond to which general concepts and vice versa. The ability for judgment is a foundation of all higher spiritual abilities, including the beautiful and sublime. Kant's aesthetics is a foundation of aesthetic theories till this very day (we will continue this discussion later). Sartre equated emotions, to a significant extent, with unconscious contents of the psyche; today this does not seem to be adequate. Jung analyzed conscious and unconscious aspects of emotions. He emphasized undifferentiated status of primitive fused emotion-concept-behavior psychic states in everyday functioning and their role in neuroses. He also emphasized rational aspect of conscious differentiated emotions. Ortony explains emotions in terms of knowledge representations and emphasizes abductive logic as a mechanism of inferencing other people's emotions. LeDoux analyses neural structures and pathways involved in emotional processing, especially in fear. Griffiths considers basic emotions and their evolutionary development within social interactions. According to Damasio, emotions are primarily bodily perceptions, and feelings of emotions in the brain invoke "bodily markers". Grossberg and Levine consider emotions as neural signals that relate instinctual and conceptual brain centers. In processes of perception and cognition, emotions evaluate concept-models of objects and situations for satisfaction or dissatisfaction of instinctual needs. In Section 6, I discuss relationships of these various theories of emotions to mathematical descriptions in the next section, here I just mention that this mathematical description closely corresponds to ideas of Kant, Jung, Grossberg and Levine. Ideas of Sartre and Damasio were not elaborated mathematically.

Behavior is comprised of many mechanisms. Behavior is controlled by the endocrine system and the nervous system. The complexity of the behavior of an organism is related to the complexity of its nervous system. In this review I refer only to neurally controlled behavior; it involves mechanisms of negative feedback (e.g., when reaching an object with a hand) and positive feedback (e.g. when making a decision). The first does not reach consciousness, the second is potentially available to consciousness [9].

Even this very cursory review of basic notions illustrates that they are far from being crystal clear; some notions may seem to contradict others. Below I summarize and simplify this discussion of basic mechanisms of the mind and relate them to mathematical discussions in the next section. This summarization and simplification of the huge body of discussions ongoing for millennia is inspired by trying to find unifying themes in commonsense understanding and technical writings from ancient philosophers to today's research in multiple disciplines.

Explaining basic mind mechanisms, let me repeat, requires no mysterious assumptions; each mechanism can be described mathematically. Among the mind's cognitive mechanisms, the most directly accessible to consciousness are concepts. Concepts are like internal models of the objects and situations in the world; this analogy is quite literal, e.g., during visual perception of an object, a concept-model in our memory projects an image onto the visual cortex, which is matched there to an image, projected from retina (this simplified description will be refined later).

Concepts serve for satisfaction of the basic instincts, which have emerged as survival mechanisms long before concepts. Inborn, less adaptive, unconscious, and more automatic functioning often is referred to as instinctual. This lumping together of various mechanisms is inappropriate for the development of physical theory, requiring mathematical description of the mind mechanisms. Grossberg and Levine [10] separated instincts as internal sensors indicating the basic needs, from "instinctual behavior", which should be described by appropriate mechanisms. Accordingly, I use word "instincts" to describe mechanisms of internal sensors: for example, when a sugar level in blood goes below a certain level an instinct "tells us" to eat. Such separation of instinct as "internal sensor" from "instinctual behavior" helps explaining many cognitive functions.

How do we know about instinctual needs? We do not hear instinctual pronouncements or read dials of instinctual sensors. Instincts are connected to cognition and behavior by emotions. Whereas in colloquial usage, emotions are often understood as facial expressions, higher voice pitch, exaggerated gesticulation, these are outward signs of emotions, serving for communication. A more fundamental role of emotions within the mind system is that emotional signals evaluate concepts for the purpose of instinct satisfaction [10]. This emotional mechanism is crucial for breaking out of the "vicious circle" of combinatorial complexity.

An inevitable conclusion from a mathematical analysis: humans and higher animals have a special instinct responsible for cognition. Let me emphasize, this is not an abstract mathematical theorem, but a conclusion from the basic knowledge of the mind operations as described in thousands of publications. Clearly, humans and animals engage into exploratory behavior, even when basic bodily needs, like eating, are satisfied. Biologists and psychologists discussed various aspects of this behavior. Harry Harlow discovered that monkeys as well as humans have the drive for positive stimulation, regardless of satisfaction of drives such as hunger [48]; David Berlyne discussed curiosity in this regard [49]; Leon Festinger, introduced the notion of cognitive dissonance and described many experiments on the drive of humans to reduce dissonance [50]. Until recently, however, it was not mentioned among "basic instincts" on a par with instincts for food and procreation. The reasons were that it was difficult to define, and that its fundamental nature was not obvious. The fundamental nature of this mechanism is related to the fact that our knowledge always has to be modified to fit the current situations. One rarely sees exactly the same object: illumination, angles, surrounding objects are usually different; therefore, adaptation–learning is required. A mathematical formulation of the mind mechanisms makes obvious the fundamental nature of our desire for knowledge. In fact virtually all learning and adaptive algorithms (tens of thousands of publications) maximize correspondence between the algorithm internal structure (knowledge in a wide sense) and objects of recognition. Concept-models that our mind uses for understanding the world are in a constant need of adaptation. Knowledge is not just a static state; it is in a constant process of adaptation and learning. Without adaptation of concept-models we will not be able to understand the ever-changing surrounding world. We will not be able to orient ourselves or satisfy any of the bodily needs. Therefore, we have an inborn need, a drive, an instinct to improve our knowledge. I call it *the knowledge instinct*. Mathematically it is described as a maximization of a similarity measure between concept-models and the world (as it is sensed by sensory organs; also the very sensing is usually adapted and shaped during perception).

Emotions evaluating satisfaction or dissatisfaction of the knowledge instinct are not directly related to bodily needs. Therefore, they are "spiritual" or aesthetic emotions. These aesthetic emotions are not peculiar to perception of art; they are inseparable from every act of perception and cognition. Conceptual–emotional understanding of the world results in actions in the outside world or within the mind. In this review we only discuss an internal behavior within the mind, the behavior of learning and understanding the world; the behavior of increasing knowledge. A mathematical theory of conceptual–emotional recognition and understanding is presented in the next section. Later, it is related to intuition, imagination, conscious, and unconscious; to an ability of the mind to think, to operate with symbols and signs. The mind involves a hierarchy of multiple levels of concept-models, from simple perceptual elements (like edges, or moving dots), to concept-models of objects, to relationships among objects, to complex scenes, and up the hierarchy ... toward the concept-models of the meaning of life and purpose of our existence. Hence the tremendous complexity of the mind, yet relatively few basic principles of the mind organization go a long way explaining this system.

I would like to add a side comment. In neural, cognitive, and psychological literature about the mind and brain, one often encounters a statement that the brain is a kludge, a nonelegant, nonoptimal design, a concoction of modules that appeared in evolution first for one purpose, then were used for a different purpose, etc. [51–54]. These statements are made usually by nonmathematicians, whose ideas about mathematical optimality and elegance are at best naive (we discussed that in this line of research many considered formal logic as the peak of optimality and elegance, already after Gödel proved its mathematical inconsistency). Mathematical analysis of evolution demonstrates just the opposite [55], there was more than enough information for evolution to attain optimality. The mind is often optimal [56]. Among those preaching nonoptimality of the brain and mind, no one produced a computer program working better or more optimally than the mind. Therefore, it is reasonable to consider mathematically optimal methods for modeling the mind (of course, mathematical descriptions have been compared with experimental data, and continuous comparison of theory and experiments is an essential part of physics).

5. Modeling field theory (MFT)

Modeling field theory mathematically implements the mechanisms of the mind discussed above. It is a multi-level, hetero-hierarchical system [8]. The mind is not a strict hierarchy; there are multiple feedback connections among adjacent levels, hence the term hetero-hierarchy. At each level in MFT there are concept-models encapsulating the mind's knowledge; they generate so-called top–down neural signals, interacting with input, bottom–up signals. These interactions are governed by the knowledge instinct, which drives concept-model learning, adaptation, and formation of new concept-models for better correspondence to the input signals.

This section describes a basic mechanism of interaction between two adjacent hierarchical levels of bottom–up and top–down signals (fields of neural activation; in this aspect MFT follows [57]); sometimes, it will be more convenient to talk about these two signal-levels as an input to and output from a (single) processing-level. At each level, output signals are concepts recognized in (or formed from) input signals. Input signals are associated with (or recognized, or grouped into) concepts according to the models and the knowledge instinct at this level. This general structure of MFT corresponds to our knowledge of neural structures in the brain; still, in this review we do not map mathematical mechanisms in all their details to specific neurons or synaptic connections. The knowledge instinct is described mathematically as maximization of a similarity measure. In the process of learning and understanding input signals, models are adapted for better representation of the input signals so that similarity between the models and signals increases. This increase in similarity satisfies the knowledge instinct and is felt as aesthetic emotions.

5.1. The knowledge instinct

At a particular hierarchical level, we enumerate neurons by index $n = 1, \ldots, N$. These neurons receive bottom–up input signals, $\mathbf{X}(n)$, from lower levels in the processing hierarchy. $\mathbf{X}(n)$ is a field of bottom–up neuronal synapse activations, coming from neurons at a lower level. Each neuron has a number of synapses; for generality, we describe each neuron activation as a set of numbers, $\mathbf{X}(n) = \{X_d(n),\ d = 1, \ldots, D\}$. Top–down, or priming signals to these neurons are sent by concept-models, $\mathbf{M}_h(\mathbf{S}_h, n)$; we enumerate models by index $h = 1, \ldots, H$. Each model is characterized by its parameters, $\mathbf{S}_h$; in the neuron structure of the brain they are encoded by strength of synaptic connections, mathematically, we describe them as a set of numbers, $\mathbf{S}_h = \{S_h^a,\ a = 1, \ldots, A\}$. Models *represent* signals in the following way. Say, signal $\mathbf{X}(n)$, is coming from sensory neurons activated by object h, characterized by parameters $\mathbf{S}_h$. These parameters may include position, orientation, or lighting of an object h. Model $\mathbf{M}_h(\mathbf{S}_h, n)$ predicts a value $\mathbf{X}(n)$ of a signal at neuron n. For example, during visual perception, a neuron n in the visual cortex receives a signal $\mathbf{X}(n)$ from retina and a priming signal $\mathbf{M}_h(\mathbf{S}_h, n)$ from an object–concept-model h. A neuron n is activated if both bottom–up signal from lower-level-input and top–down priming signal are strong. Various models compete for evidence in the bottom–up signals, while adapting their parameters for better match as described below. This is a simplified description of perception. The most benign everyday visual perception uses many levels from retina to object perception. The MFT premise is that the same laws describe the basic interaction dynamics at each level. Perception of minute features, or everyday objects, or cognition of complex abstract concepts is due to the same mechanism described below. Perception and cognition involve models and learning. In perception, models correspond to objects; in cognition models correspond to relationships and situations.

Learning is an essential part of perception and cognition, and it is driven by the knowledge instinct. It increases a similarity measure between the sets of models and signals, $L(\{\mathbf{X}\},\{\mathbf{M}\})$. The similarity measure is a function of model parameters and associations between the input bottom–up signals and top–down, concept-model signals. For concreteness I refer here to an object perception using a simplified terminology, as if perception of objects in retinal signals occurs in a single level.

In constructing a mathematical description of the similarity measure, it is important to acknowledge two principles (which are almost obvious). First, the visual field content is unknown before perception occurred and second, it may contain any of a number of objects. Important information could be contained in any bottom–up signal; therefore, the similarity measure is constructed so that it accounts for all bottom–up signals, $\mathbf{X}(n)$,

$$L\big(\{\mathbf{X}\},\{\mathbf{M}\}\big)=\prod_{n\in N} l\big(\mathbf{X}(n)\big). \tag{1}$$

This expression contains a product of partial similarities, $l(\mathbf{X}(n))$, over all bottom–up signals; therefore it forces the mind to account for every signal (even if one term in the product is zero, the product is zero, the similarity is low and the knowledge instinct is not satisfied); this is a reflection of the first principle. Second, before perception occurs, the mind does not know which object gave rise to a signal from a particular retinal neuron. Therefore a partial similarity measure is constructed so that it treats each model as an alternative (a sum over models) for each input neuron signal. Its constituent elements are conditional partial similarities between signal $\mathbf{X}(n)$ and model $\mathbf{M}_h$, $l(\mathbf{X}(n)|h)$. This measure is "conditional" on object h being present, [58] therefore, when combining these quantities into the overall similarity measure, L, they are multiplied by $r(h)$, which represent a probabilistic measure of object h actually being present. Combining these elements with the two principles noted above, a similarity measure is constructed as follows [59]:

$$L\big(\{\mathbf{X}\},\{\mathbf{M}\}\big)=\prod_{n\in N}\sum_{h\in H} r(h)\, l\big(\mathbf{X}(n)|h\big). \tag{2}$$

The structure of (2) follows standard principles of the probability theory: a summation is taken over alternatives, h, and various pieces of evidence, n, are multiplied. This expression is not necessarily a probability, but it has a probabilistic structure. If learning is successful, it approximates probabilistic description and leads to near-optimal Bayesian decisions. The name "conditional partial similarity" for $l(\mathbf{X}(n)|h)$ (or simply $l(n|h)$) follows the probabilistic terminology. If learning is successful, $l(n|h)$ becomes a conditional probability density function, a probabilistic measure that signal in neuron n originated from object h. Then L is a total likelihood of observing signals $\{\mathbf{X}(n)\}$ coming from objects described by models $\{\mathbf{M}_h\}$. Coefficients $r(h)$, called priors in probability theory, contain preliminary biases or expectations, expected objects h have relatively high $r(h)$ values; their true values are usually unknown and should be learned, like other parameters $\mathbf{S}_h$.

We note that in probability theory, a product of probabilities usually assumes that evidence is independent. Expression (2) contains a product over n, but it does not assume independence among various signals $\mathbf{X}(n)$. There is a dependence among signals due to models: each model $\mathbf{M}_h(\mathbf{S}_h,n)$ predicts expected signal values in many neurons n.

During the learning process, concept-models are constantly modified. In this review we consider a case when functional forms of models, $\mathbf{M}_h(\mathbf{S}_h,n)$, are all fixed and learning–adaptation involves only model parameters, $\mathbf{S}_h$. More complicated structural learning of models is considered in [60,61]. From time to time a system forms a new concept, while retaining an old one as well; alternatively, old concepts are sometimes merged or eliminated. This requires a modification of the similarity measure (2); the reason is that more models always result in a better fit between the models and data. This is a well-known problem, it is addressed by reducing similarity (2) using a "skeptic penalty function", $p(N,M)$ that grows with the number of models M, and this growth is steeper for a smaller amount of data N. For example, an asymptotically unbiased maximum likelihood estimation leads to multiplicative $p(N,M)=\exp(-N_{\text{par}}/2)$, where N_{par} is a total number of adaptive parameters in all models (this penalty function is known as Akaike Information Criterion, see [8] for further discussion and references).

5.2. Dynamic logic

The learning process consists in estimating model parameters **S** and associating signals with concepts by maximizing the similarity (2). Note, all possible combinations of signals and models are accounted for in expression (2). This can be seen by expanding a sum in (2), and multiplying all the terms; it would result in H^N items, a huge number. This is the number of combinations between all signals (N) and all models (H). Here is the source of CC of many algorithms used in the past. For example, multiple hypothesis testing algorithms attempts to maximize similarity L over model parameters and associations between signals and models, in two steps. First it takes one of the H^N items, which is one particular association between signals and models; and maximizes it over model parameters. Second, the largest item is selected (that is the best association for the best set of parameters). Such a program inevitably faces a wall of CC, the number of computations on the order of H^N.

Modeling field theory solves this problem by using dynamic logic [8,62]. An important aspect of dynamic logic is matching vagueness or fuzziness of similarity measures to the uncertainty of models. Initially, parameter values are not known, and uncertainty of models is high; so is the fuzziness of the similarity measures. In the process of learning, models become more accurate, and the similarity measure more crisp, the value of the similarity increases. This is the mechanism of dynamic logic.

Mathematically it is described as follows. First, assign any values to unknown parameters, $\{\mathbf{S}_h\}$. Then, compute association variables $f(h|n)$,

$$f(h|n) = r(h)l\big(\mathbf{X}(n)|h\big) \Big/ \sum_{h' \in H} r(h')l\big(\mathbf{X}(n)|h'\big). \tag{3}$$

Eq. (3) looks like the Bayes formula for a posteriori probabilities; if $l(n|h)$ in the result of learning become conditional likelihoods, $f(h|n)$ become Bayesian probabilities for signal n originating from object h. The dynamic logic of the Modeling Fields (MF) is defined as follows:

$$\mathrm{d}f(h|n)/\mathrm{d}t = f(h|n) \sum_{h' \in H} \big[\delta_{hh'} - f(h'|n)\big] \cdot \big[\partial \ln l(n|h')/\partial \mathbf{M}_{h'}\big] \partial \mathbf{M}_{h'}/\partial \mathbf{S}_{h'} \cdot \mathrm{d}\mathbf{S}_{h'}/\mathrm{d}t, \tag{4}$$

$$\mathrm{d}\mathbf{S}_h/\mathrm{d}t = \sum_{n \in N} f(h|n)\big[\partial \ln l(n|h)/\partial \mathbf{M}_h\big] \partial \mathbf{M}_h/\partial \mathbf{S}_h, \tag{5}$$

here

$$\delta_{hh'} \quad \text{is} \quad \begin{cases} 1 & \text{if } h = h', \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

Parameter t is the time of the internal dynamics of the MF system (like a number of internal iterations). Gaussian-shape functions can often be used for conditional partial similarities,

$$l(n|h) = G\big(\mathbf{X}(n)|\mathbf{M}_h(\mathbf{S}_h, n), \mathbf{C}_h\big). \tag{7}$$

Here G is a Gaussian function with mean $\mathbf{M}_h$ and covariance matrix $\mathbf{C}_h$. Note, a "Gaussian assumption" is often used in statistics; it assumes that signal distribution is Gaussian. This is not the case in (7): here signal is not assumed to be Gaussian. Eq. (7) is valid if *deviations* between the model **M** and signal **X** are Gaussian; these deviations usually are due to many random causes and, therefore, Gaussian. If they are not Gaussian, appropriate functions could be used. If there is no information about functional shapes of conditional partial similarities, still (7) is a good choice, it is not a limiting assumption: a weighted sum of Gaussians in (2) can approximate any positive function, like similarity.

Covariance matrices, $\mathbf{C}_h$, in (7) are estimated like other unknown parameters. Their initial values should be large, corresponding to uncertainty in knowledge of models, $\mathbf{M}_h$. As parameter values and models improve, covariances are reduced to intrinsic differences between models and signals (due to sensor errors, or model inaccuracies). As covariances get smaller, similarities get crisper, closer to delta-functions; association variables (3) get closer to crisp $\{0, 1\}$ values, and dynamic logic solutions converge to crisp logic. This process of concurrent parameter improvement and convergence of similarity to a crisp logical function is an essential part of dynamic logic. This is the mechanism of dynamic logic combining fuzzy and crisp logic.

The dynamic evolution of fuzziness from large to small is the reason for the name "dynamic logic". Mathematically, this mechanism helps avoiding local maxima during convergence [8], and psychologically it explains many properties of the mind, as discussed in Section 6 [63]. Whichever functional shapes are used for conditional partial similarities, they ought to allow for this process of matched convergence in parameter values and similarity crispness.

The following theorem was proved [8].

Theorem. *Eqs. (3)–(6) define a convergent dynamic MF system with stationary states defined by* $\max_{\{\mathbf{S}h\}} L$.

It follows that the stationary states of an MF system are the maximum similarity states satisfying the knowledge instinct. When partial similarities are specified as probability density functions (pdf), or likelihoods, the stationary values of parameters $\{\mathbf{S}_h\}$ are asymptotically unbiased and efficient estimates of these parameters [64]. A computational complexity of the MF method is linear in N.

In plain English, this means that dynamic logic is a convergent process. It converges to the maximum of similarity, and therefore satisfies the knowledge instinct. Several aspects of MFT convergence are discussed below (in Sections 5.3, 5.4, and 6.1). If likelihood is used as similarity, parameter values are estimated efficiently (that is, in most cases, parameters cannot be better learned using any other procedure). Moreover, as a part of the above theorem, it is proven that the similarity measure increases at each iteration. The psychological interpretation is that the knowledge instinct is satisfied at each step: a modeling field system with dynamic logic *enjoys* learning.

5.3. Example of dynamic logic operations

Finding patterns below noise can be an exceedingly complex problem. If an exact pattern shape is not known and depends on unknown parameters, these parameters should be found by fitting the pattern model to the data. However, when the locations and orientations of patterns are not known, it is not clear which subset of the data points should be selected for fitting. A standard approach for solving this kind of problem, which has already been discussed, is multiple hypothesis testing [27]. Here, since all combinations of subsets and models are exhaustively searched, it faces the problem of combinatorial complexity. In the current example, we are looking for "smile" and "frown" patterns in noise shown in Fig. 1(a) without noise, and in Fig. 1(b) with noise, as actually measured. Each pattern is characterized by a 3-parameter parabolic shape. The image size in this example is 100×100 points, and the true number of patterns is 3, which is not known. Therefore, at least 4 patterns should be fit to the data, to decide that 3 patterns fit best. Fitting $4 \times 3 = 12$ parameters to 100×100 grid by a brute-force testing would take about 10^{32} to 10^{40} operations, a prohibitive computational complexity.

To apply MFT and dynamic logic to this problem one needs to develop parametric adaptive models of expected patterns. The models and conditional partial similarities for this case are described in details in [65]: a uniform model for noise, Gaussian blobs for highly-fuzzy, poorly resolved patterns, and parabolic models for "smiles" and "frowns". The number of computer operations in this example was about 10^{10}. Thus, a problem that was not solvable due to CC becomes solvable using dynamic logic.

During an adaptation process, initial fuzzy and uncertain models are associated with structures in the input signals, and fuzzy models become more definite and crisp with successive iterations. The type, shape, and number, of models are selected so that the internal representation within the system is similar to input signals: the MF concept-models represent structure-objects in the signals. The figure below illustrates operations of dynamic logic. In Fig. 1(a) true "smile" and "frown" patterns are shown without noise; (b) actual image available for recognition (signal is below noise, signal-to-noise ratio is between $-2\mathrm{d}B$ and $-0.7\mathrm{d}B$); (c) an initial fuzzy model, a large fuzziness corresponds to uncertainty of knowledge; (d) through (h) show improved models at various iteration stages (total of 22 iterations). Every five iterations the algorithm tried to increase or decrease the number of pattern-models. Between iterations (d) and (e) the algorithm decided, that it needs three Gaussian models for the "best" fit. There are several types of models: one uniform model describing noise (it is not shown) and a variable number of blob models and parabolic models, which number, location, and curvature are estimated from the data. Until about stage (g) the algorithm used simple blob models, at (g) and beyond, the algorithm decided that it needs more complex parabolic models to describe the data. Iterations stopped at (h), when similarity stopped increasing.

5.4. MFT hierarchical organization

Above, we described a single processing level in a hierarchical MFT system. At each level of a hierarchy there are input signals from lower levels, models, similarity measures (2), emotions, which are changes in similarity (2),

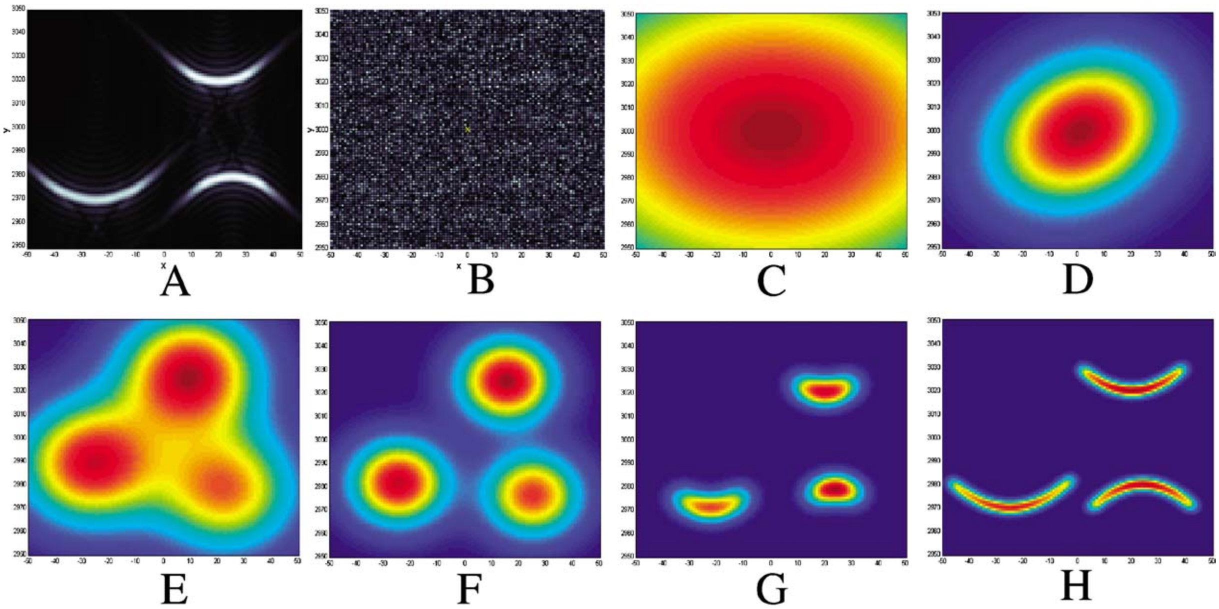

Fig. 1. Finding "smile" and "frown" patterns in noise, an example of dynamic logic operation: (a) true "smile" and "frown" patterns are shown without noise; (b) actual image available for recognition (signal is below noise, signal-to-noise ratio is between $-2\mathrm{d}B$ and $-0.7\mathrm{d}B$); (c) an initial fuzzy blob-model, the fuzziness corresponds to uncertainty of knowledge; (d) through (h) show improved models at various iteration stages (total of 22 iterations). Between stages (d) and (e) the algorithm tried to fit the data with more than one model and decided, that it needs three blob-models to "understand" the content of the data. There are several types of models: one uniform model describing noise (it is not shown) and a variable number of blob-models and parabolic models, which number, location, and curvature are estimated from the data. Until about stage (g) the algorithm "thought" in terms of simple blob models, at (g) and beyond, the algorithm decided that it needs more complex parabolic models to describe the data. Iterations stopped at (h), when similarity (2) stopped increasing. This example is discussed in more details in [65].

and actions; actions include adaptation, behavior satisfying the knowledge instinct—maximization of similarity, Eqs. (3)–(6). An input to each level is a set of signals $\mathbf{X}(n)$, or in neural terminology, an input field of neuronal activations. The result of signal processing at a given level are activated models, or concepts h recognized in the input signals n; these models along with the corresponding instinctual signals and emotions may activate behavioral models and generate behavior at this level.

The activated models initiate other actions. They serve as input signals to the next processing level, where more general concept-models are recognized or created. Output signals from a given level, serving as input to the next level, could be model activation signals, a_h, defined as

$$a_h = \sum_{n\in N} f(h|n). \tag{8}$$

Alternatively, output signals may include model parameters. The hierarchical MF system is illustrated in Fig. 2. Within the hierarchy of the mind, each concept-model finds its "mental" meaning and purpose at a higher level (in addition to other purposes). For example, consider a concept-model "chair". It has a "behavioral" purpose of initiating sitting behavior (if sitting is required by the body), this is the "bodily" purpose at the same hierarchical level. In addition, it has a "purely mental" purpose at a higher level in the hierarchy, a purpose of helping to recognize a more general concept, say of a "concert hall", which model contains rows of chairs.

Models at higher levels in the hierarchy are more general than models at lower levels. For example, at the very bottom of the hierarchy, if we consider vision system, models correspond (roughly speaking) to retinal ganglion cells and perform similar functions; they detect simple features in the visual field; at higher levels, models correspond to functions performed at V1 and higher up in the visual cortex, that is detection of more complex features, such as contrast edges, their directions, elementary moves, etc. Visual hierarchical structures and models are studied in details [9,66], these models can be used in MFT. At still higher cognitive levels, models correspond to objects, to relationships among objects, to situations, and relationships among situations, etc. [8]. Still higher up are even more general models of complex cultural notions and relationships, like family, love, friendship, and abstract concepts, like law, rationality,

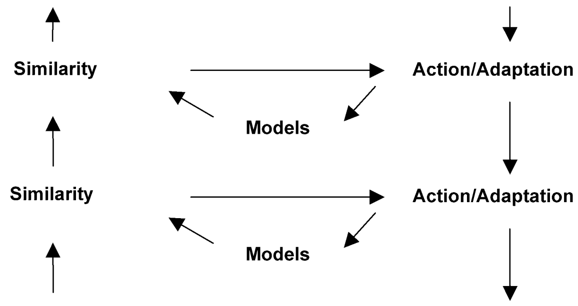

Fig. 2. Hierarchical MF system. At each level of a hierarchy there are models, similarity measures, and actions (including adaptation, maximizing the knowledge instinct—similarity). High levels of partial similarity measures correspond to concepts recognized at a given level. Concept activations are output signals at this level and they become input signals to the next level, propagating knowledge up the hierarchy.

etc. Contents of these models correspond to cultural wealth of knowledge, including writings of Shakespeare and Tolstoy; mechanisms of the development of these models are reviewed in the next section. At the top of the hierarchy of the mind, according to Kantian analysis [67], are models of the meaning and purpose of our existence, unifying our knowledge, and the corresponding behavioral models aimed at achieving this meaning.

From time to time, as discussed, a system forms a new concept or eliminates an old one. Many pattern recognition algorithms and neural networks luck this important ability of the mind. It can be modeled mathematically in several ways; adaptive resonance theory (ART) uses vigilance threshold, which is compared to a similarity measure [57]. A somewhat different mechanism of MFT works as follows. At every level, the system always keeps a reserve of vague (fuzzy) inactive concept-models (with large covariance, **C**, Eq. (7)). They are inactive in that their parameters are not adapted to the data, therefore their similarities to signals are low. Yet, because of a large fuzziness (covariance) the similarities are not exactly zero. When a new signal does not fit well into any of the active models, its similarities to inactive models automatically increase (because first, every piece of data is accounted for see Ref. [58], and second, inactive models are vague-fuzzy and potentially can "grab" every signal that does not fit into more specific, less fuzzy, active models). When the activation signal a_h of Eq. (8) for an inactive model, h, exceeds a certain threshold, the model is activated. Similarly, when an activation signal for a particular model falls below a threshold, the model is deactivated. Thresholds for activation and deactivation are set usually based on information existing at a higher hierarchical level (prior information, system resources, numbers of activated models of various types, etc.). Activation signals for active models at a particular level $\{a_h\}$ form a "neuronal field", which serve as input signals to the next level, where more abstract and more general concepts are formed, and so on along the hierarchy toward higher models of meaning and purpose.

6. Symbolic culture

6.1. Symbols in cognitive linguistics, philosophy, and computational intelligence

"Symbol is the most misused word in our culture" [68]. We use this word in trivial cases referring to traffic signs, and in the most profound cases to cultural and religious symbols. Charles Peirce considered symbols to be a particular type of signs [69]. He concentrated on the process of sign interpretation, which he conceived as a triadic relationship of sign, object, and interpretant. Interpretant is similar to what we call today a representation of the object in the mind. Peircian approach, however, was inconsistent. He classified signs into symbols, indexes, and icons. Icons have meanings due to resemblance to the signified (objects, situations, etc.), indexes have meanings by direct connection to the signified, and *symbols* have meaning due to arbitrary conventional agreements. However, identifying "resemblance" turned out a complex problem; pattern recognition algorithms based on simple resemblances do not work. Similarly problematic was an idea of "arbitrary" symbols, as we discuss later.

When scientists attempted to understand symbols in the development of semiotics, the two functions, *understanding language* and *understanding world*, were often mixed up, as if cognition was provided by language. This tendency was strengthened by considering logic to be the mechanism of both, language and cognition. According to Bertrand Russell, language is equivalent to axiomatic logic, "[a word-name is] merely to indicate what we are speaking about;

[it] is no part of the fact asserted... it is merely part of the symbolism by which we express our thought" [70]. Logical positivism centered on "the elimination of metaphysics through the logical analysis of language"—according to Rudolf Carnap [71] logic was sufficient for the analysis of language. We have already analyzed roots of this belief in logic.

Similar understanding of relationships among symbol, language, logic, and the mind can be traced in semiotics of Ferdinand Saussure and in structuralism. A simplistic idea that words are labels for objects falls apart as soon as we consider words for abstract ideas, say, "rational". Saussure tried to resolve this problem by saying that "the linguistic sign does not unite a thing and a name, but a concept and a sound image" [72]. Here, both aspects of the sign exist in the mind and the real world is taking a back sit. Like in formal logic, relationships between mathematical objects and the world are arbitrary. Similarly, Saussure emphasized the arbitrariness of the sign, relationships between words and objects in the world are arbitrary conventions.

This idea later heavily influenced Structuralism, and evolved into arbitrariness of communication codes in general. Since communication codes contain cultural values, some concluded that cultural values are arbitrary. "There may be an objective, empiricist reality out there, but there is no universal, objective way of perceiving and making sense of it. What passes for reality in any culture is the product of the culture's codes, so "reality" is always already encoded, it is never "raw" [73]. This circle of ideas served as a platform for Jacques Derrida's attacks on structuralism [74]. Since any statement is based on some cultural structures and values, it can be dismissed as having no objective validity, as "arbitrary" or as "local". Here we review attempts [60,61] to give answer to this question: How is it possible to have anything of truth and value? How our mind constructs symbols, which have psychological values and are not reducible to arbitrary signs.

A path toward such understanding was not straight. The first move towards mathematical models of language and cognition was in opposite directions. It took a while for science, first, to appreciate that language and cognition were not one and the same, and later to realize that their relationships require scientific elucidation. In the second half of the 20th century Noam Chomsky reoriented linguistics toward studies of innate mind mechanisms [75]. "Nativists" (as followers of Chomsky are called today) emphasized that mechanisms of language were separate from the rest of cognition. In the 1960s and 1970s they used mathematics of logical rules, similar to artificial intelligence. In the 1980s Chomsky proposed a new mathematical paradigm in linguistics, rules and parameters [76]. This was similar to model-based systems emerging in mathematical studies of cognition. In 1990s, Chomsky's minimalist program called for simplifying rule structure of the mind mechanism of language [77]. It moved language closer to other mind mechanisms, closer to cognition, but stopped at an interface between language and cognition. Nativist linguists, following Chomsky, emphasized that cognition and language abilities are different: they are located in different brain areas and they might have emerged along separate paths in evolution [78]. Learning language models is driven by the language instinct, separate from the knowledge instinct. Chomsky's linguistics still assumed that meanings appear independently from language, and a mix-up of signs and symbols continued; motivational forces of symbols were ignored.

Cognitive linguistics emerged in the 1970s to address some of these limitations of the nativist approach. "Cognitivists" wanted to unify language with cognition and explain creation of meanings. They were looking for simpler innate structures than those postulated by nativists. These simpler structures would be sufficient, scientists thought, because they will combine language and cognition, combine innate structures with learning from experience (to a much larger extent than nativists postulated). Cognitivists gradually moved away from heavy logical bias of the previous structuralist thinking, which could be characterized by

> "the meaning of a word can be exhaustively decomposed into finite set of conditions ... necessary and sufficient ..." [79]

George Lakoff emphasized that abstract concepts used by the mind for understanding the world have metaphorical structure [80]. Metaphors were not just poetic tools, but an important mechanism of the mind for creating new abstract meanings. Lakoff' analysis brought this cultural knowledge, advanced by Fyodor Dostoevsky and Friedrich Nietzsche, within the mainstream of science. There was still a big gap between Lakoff' analysis of metaphors [81] on one hand and neural and mathematical mechanisms on the other. The "Metaphors we live by" is a metaphorical book in that it begs the question: Who is that homunculus in the mind, interpreting the metaphorical theater of the mind? What are the mechanisms of metaphorical thinking?

In works of Jackendoff [79], Langacker [82], Talmy [83], and other cognitive linguists [84] it was recognized that old divisions dominating linguistics were insufficient. Dichotomies of meanings (semantic-pragmatic), dichotomies of hierarchical structures (superordinate-subordinate) were limiting scientific discourse and had to be overcome. Consider the following opinions on meaning creation:

> "in a hierarchical structure of meaning determination the superordinate concept is a necessary condition for the subordinate one ... COLOR is a necessary condition for determining the meaning or RED" [79]

> "The base of predication is nothing more than ... domains which the prediction actually invokes and requires" [82]

These examples illustrate attempts to overcome old dichotomies and at the same time difficulties encountered along this path. Both examples are influenced by logical bias. Attempts to implement mathematically mechanisms assumed by these examples would lead to combinatorial complexity. To put it jovially, problems of meaning and hierarchy still reminded the old question about the chicken and the egg, what came first? If superordinate concepts come before subordinate ones, where do they come from? Are we born with the concept COLOR in our minds? If predictions invoke domains, where do domains come from? These complex questions with millennial pedigrees are answered mathematically in the following sections. Here I give a brief psychological preview of the answer, accounting for contemporary development in dynamic logic, neurobiology, and language evolution. Hierarchy and meaning emerge jointly with cognition and language. In processes of evolution and individual learning, superordinate concepts (COLOR) are vaguer, less specific, and less conscious than subordinate ones (RED). RED can be vividly perceived, but COLOR cannot be perceived. RED can be perceived by animals. But, the concept COLOR can only emerge in the human mind, due to joint operation of language and cognition.

Elman [85] emphasized that cognitively driven, use-based language acquisition is possible, using connectionist (neural network) mechanisms. The main argument was that the innate mechanisms of connectionist architectures can be much simpler than logical rules postulated by nativists. But what exactly is "simpler"? Elman himself emphasized this other side of the story. The connectionist neural network is not an off-the-shelf product, but SNR neural network carefully designed for language acquisition (Elman [86]). Moreover, SNR does not perform "general" language acquisition, but a specific type of learning it was designed for. Elman emphasized a hard learned lesson: "there is no ... general purpose learning algorithm that works equally well across domains" [85, p. 1]. We already discussed this in Sections 2 and 3.

Jackendoff in his recent research [98] concentrated on unifying language and cognition. He developed detailed models for such unification; however, his logical structures face combinatorial complexity.

We have, however, reasons to believe that the mind *does* work, and that there are the first principles of the mind organization [87] leading to an ability for creating symbols. SNR neural network cannot be an example for such a general principle: according to analysis in previous sections, SNR will face combinatorial complexity, when exposed to complex learning. It will not scale up to real human brain. Elman [88] is among the first to admit this. Still, SNR can be used to elucidate the general principles. Among such principles is evolution of abstract notions from vague and fuzzy toward specific and concrete (Elman [85, p. 14]; Olguin and Tomasello [89]; Tomasello and Olguin [90]). Dynamic logic systematically utilizes this principle. We already addressed another important principle of the mind organization discussed recently by Nolfi et al. [91], learning is motivated by internal drives. There is an important difference, however, between Elman [88] discussion of nonspecific emergence and the purposeful emergence mechanisms of the instinct for knowledge.

Michael Tomasello [92,93] suggests that the first principle of the human mind organization, the most important mechanism of the human brain required to learn language is not language specific, but more broadly cultural and social. It is our ability to perceive other people as intentional agents. We understand that other people have intentions and plans to achieve them, we can figure out what these intentions and plans are. This is the foundation for our entire symbolic culture. Mathematical or neural mechanisms of this ability are not known. In the following sections I describe a mathematical theory of joint learning of cognition and language. Its most important premise is that we are born with an innate drive, an instinct for knowledge. It determines the purposiveness of our existence, our higher mental abilities, and our ability to create symbolic culture. It is mathematically possible that a significant aspect of this drive is to acquire knowledge about other people' intentions and plans. It would be a fascinating enterprise to establish relationships between these two theories through mathematical modeling, neural research, and psychological laboratory experimentation.

Let us summarize goals and achievements of cognitive linguistics. Connectionist architectures demonstrated that complex syntax patterns can be learned without explicit rules and without explicit examples. They demonstrated elements of joint language learning and meaning creation (cognition). Still these type architectures face CC and do not scale up. Motivational forces inherent to symbols, which were recognized by Saussure and analytic psychology, made inroads into linguistics and psychology. Still, symbols and signs continue to be mixed up.

Learning cognitive models, as discussed in Section 5 is driven by the knowledge instinct. Most importantly, cognition is about objects and situations in the surrounding world, whereas mechanisms of acquiring and using language identified in cognitive linguistics are about language, not about the world. Today, we still do not know neural mechanisms combining language with cognition, nor their locations in the brain; and until recently, no mathematical mechanisms were suggested for unifying new findings in cognitive linguistics and computational intelligence. Researchers in cognition, linguistics, evolutionary linguistics, computational semiotics came to appreciation that language cannot be understood separately from thinking about the world [60,94–96]. Similar ideas are becoming accepted by Chomsky and by nativist linguists [97,98]. The next section discusses steps toward mathematical theory of symbols integrating cognition and language.

6.2. Integration of language and cognition

Integration of language and cognition in MFT [60,61] is attained by integrating cognitive and language models, so that a concept-model (in (1) through (5)) $\mathbf{M}_h$ is given by

$$\mathbf{M}_h = \left\{\mathbf{M}_h^C, \mathbf{M}_h^L\right\}. \tag{9}$$

Here $\mathbf{M}_h^C$ denotes a cognitive part of the model of an object or situation in the world (like models in example, Fig. 1), and $\mathbf{M}_h^L$ is a language part of the model. Mathematical mechanisms of integrating cognition and language require extension of MFT considered in [60]. Consider now this integrated model as the mind's mechanism of integrating language and cognition. A data stream constantly comes into the mind from all sensory perceptions; every part of this data stream is constantly evaluated and associated with models (9) according to the mechanisms of dynamic logic described in previous sections. At the beginning, the models are fuzzy; cognitive models vaguely correspond to uncertain undifferentiated sensory perceptions. Language models vaguely correspond to sounds. This is approximately a state of the mind of a newborn baby. First, models of simple perceptions differentiate; objects are distinguished in visual perception. Language sounds are differentiated from other sounds. In (9) some cognitive models become crisper than other cognitive models. Until about one year of age, perception models corresponding to simple objects become crisper at a faster rate than language models.

Gradually, models are adapted, their correspondence to specific signals improve, selectivity to language signals and nonlanguage sounds is enhanced. Language models are associated with words (sentences, etc.), and cognitive models are associated with objects and situations of perception and cognition. Between the first and second year of life the speed of adaptation of language models tremendously accelerates and overtakes learning of cognitive models. By the age of 5 or 7, a child knows tremendous number of language models (words, rules of grammar), which attained differentiated, crisp status. But it will take the rest of his life to associate them with real life situations and acquire highly differentiated crisp cognitive models.

Association between language and cognitive models occurs before any of the models attain a high degree of specificity characteristic of the grown-up conscious concepts. Language and cognition are integrated at a pre-conscious level. Certain language models evolve faster than their corresponding cognitive models and vice versa. Correspondingly, uncertainty and fuzziness of the two aspects of integrated models may significantly differ. Still, existence of crisp language models helps to identify relevant objects and situations in the world, and therefore, speeds up learning and adaptation of the corresponding cognitive models and v.v. This was suggested as a mechanism of interaction between language and cognition [60,61]. Both abilities enhance each other.

This mechanism of interaction between language and cognition was suggested for ontological development and learning in each human individual, as well as for biological evolution of the human specie, and for evolution of cultures. Few human individuals succeed in connecting the entire wealth of language models with crisp and conscious

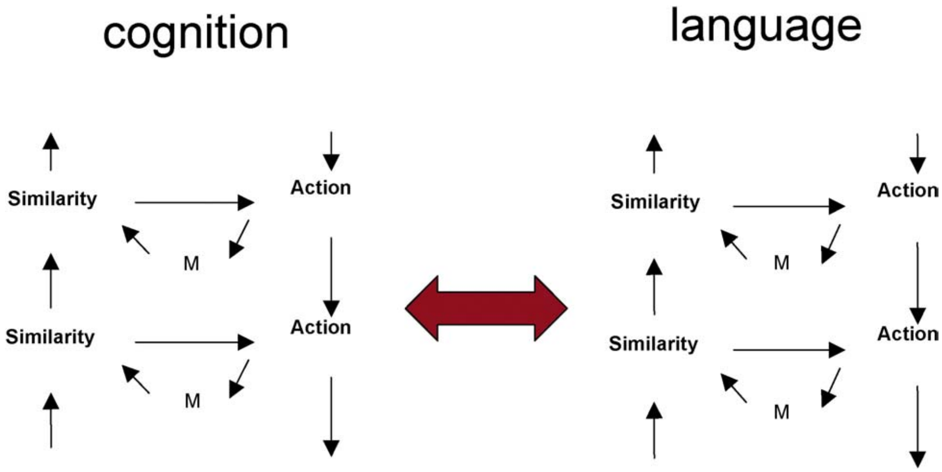

Fig. 3. Hierarchical integrated language-cognition MF system. At each level in a hierarchy there are integrated language and cognition models. Similarities are integrated as products of language and cognition similarities. Initial models are fuzzy placeholders, so integration of language and cognition is sub-conscious. Association variables depend on both language and cognitive models and signals. Therefore language model learning helps cognitive model learning and v.v. Abstract cognitive concepts are grounded in abstract language concepts.

cognitive models. Still, experience (cognitions) of the entire culture is accumulated in language. Evolution of culture is gradually accumulated in genes.

Structure (9) gives a fuzzy placeholder for a cognitive model corresponding to each language model, and v.v. It enables independent learning of language and cognitive parts of models, while enhancing each other's learning. In this way, models of each type are gradually learned, cognition helps language, and language helps cognition. Knowledge is accumulated in culture through generations.

Section 5.4 described the hierarchical MFT organization, as shown in Fig. 2. Combined with the mechanism of integrated models (9), it can integrate cognitive and language hierarchies as illustrated in Fig. 3. An amazing aspect of the human mind is that these two hierarchies are integrated in such a way that relationships among constituent models are preserved. For example, a cognitive model of a situation and the corresponding phrase model are constituted from lower-level models: objects and words. Correspondence between these objects and words in the object-word level is the same as between them, when they become constituent parts of the phrase-situation level model. And this holds true across tremendous number of the phrase-situation level models, using various combinations of the same words from the lower level. This amazing property of our mind seems so obvious, that nontrivial complexity of the required mechanism was noticed only recently [68].

Let us elaborate a bit. A dog can learn to bring shoes on command. The dog can associate shoes with a word "shoes". Does it mean dog's mind possesses models (9)? Try to teach a meaning of a word "rational" to a dog. Apparently, a dog can associate sounds with objects, which it sees in the world. A dog treats sounds just like other objects. But it does not possess a hierarchy of integrated models. In dog's mind, cognitive models are "grounded" in objects and situations in the world. But abstract concepts require grounding in other concepts, a hierarchy of concepts is required. According to [68], smartest apes after years of training, could possibly learn 2 levels of a hierarchy. Why is it so difficult? Higher levels of a hierarchy in the ape mind have no "ground". In the human mind, higher level language models are grounded in conversations with other people: Mutual understanding "assures" our mind of the reality of language hierarchy. A cognitive hierarchy is supported by a language hierarchy.

Dawkins [99] called concept-models of the mind "memes" and emphasized that model selection will overtake gene selection because models are more efficient replicators. A mathematical description of this process, as interaction of language and cognition is a subject of this review. Cognitive models that proved useful in life and evolution cannot be directly transferred to the minds of the next generation. Only language models are transferred to the next generation. This separation between cognitive models and language models can be compared to separation between phenotypes and genotypes. In some ways this comparison could be deep and inspiring, in other ways, it is superficial and wrong. Because of space limitation, we will not pursue it here.

Cognitive models created by each generation are accumulated in culture due to language. Cultural evolution selects useful models. Language accumulates cultural knowledge at all levels in a hierarchy of the mind. Due to integration of language and cognition, language provides grounding for abstract high-level cognitive models. But, this requires that every next generation connects language and cognitive models in individual minds. Every generation has to learn differentiated conscious cognitive models corresponding to the level of differentiation accumulated in language and culture. Possibly, an essential inborn difference between human and animal minds is that we possess structures similar to Eq. (9). This might be sufficient for evolution of symbolic culture.

Is it possible to prove these assertions? Current research in evolution of languages and cultures uses mathematical simulations of communities of interacting agents [96] and related statistical models [100,101]. It was shown that a compositional hierarchical language (with phrases composed from words) evolves from noncompositional languages given a structure similar to Eq. (9) under certain conditions [102]. This research was limited by an assumption of pre-existing cognitive models. Combining these results with MFT toward self-evolving integrated hierarchical cognitive-language system was initiated in [103–105].

7. Workings of the mind: Current understanding, future research, predictions and tests

I review here how the mathematical descriptions are related to a variety of the mind functions. I would like to illustrate that current understanding of the mind merits a status of the new area of physics: The mind is described mathematically, on one hand without mysticism, and on the other hand, without reductionism, in general agreement with cognitive science, psychology, and philosophy. A variety of the mind phenomena are understood from few basic principles; some long-standing controversies in psychology, philosophy, and aesthetics are resolved; and specific predictions are made that will be tested in laboratory and mathematical simulations. Among the first principles of the mind is the knowledge instinct, which is described mathematically as maximization of similarity between concept-models and the world. This principle leads to a better understanding of the mind functioning, and to solving previously unsolved problems associated with higher mind functions, including consciousness, feelings of sublime, and beauty. In addition to studying basic mechanisms of isolated minds, I address research in language and culture evolution involving simulation of interacting minds.

7.1. *Cognition and dynamics of MFT*

Eqs. (3)–(6) describe an elementary process of perception or cognition, which maximizes knowledge. Knowledge is measured by similarity between concept-models and the world. In this process a large number of model-concepts compete for incoming signals, models are modified and new ones are formed, and eventually, connections are established between signal subsets on the one hand, and model-concepts on the other. Perception refers to processes in which the input signals come from sensory organs and model-concepts correspond to objects in the surrounding world. Cognition refers to higher levels in the hierarchy where the input signals are activation signals from concepts cognized (activated) at lower levels, whereas model-concepts are more complex, abstract, and correspond to situations and relationships among lower-level concepts.

This process is described by dynamic logic. Its salient mathematical property is a correspondence between uncertainty in models and vagueness–fuzziness in associations $f(h|n)$. During perception, as long as model parameters do not correspond to actual objects, there is no match between models and signals; many models poorly match many objects, and associations remain fuzzy. Eventually, one model (h') wins a competition for a subset $\{n'\}$ of input signals $\mathbf{X}(n)$, when parameter values match object properties; $f(h'|n)$ values become close to 1 for $n \in \{n'\}$ and 0 for $n \notin \{n'\}$. Upon the convergence, the entire set of input signals $\{n\}$ is approximately divided into subsets, each associated with one model-object. Initial fuzzy concepts become crisp concepts, approximately obeying formal logic. The general mathematical laws of cognition and perception are similar.

7.2. *Elementary thought-process*

Thought-processes or thinking involves a number of sub-processes and attributes, including internal representations and their manipulation, attention, memory, concept formation, knowledge, generalization, recognition, understanding, meaning, prediction, imagination, intuition, emotion, decisions, reasoning, goals, behavior, conscious and unconscious [8,9,12]. A "minimal" subset of these processes has to involve mechanisms for afferent and efferent signals [9], in other words, bottom–up and top–down signals coming from outside (external sensor signals) and from inside (internal representation signals). According to Carpenter and Grossberg [57] every recognition and concept formation process involves a "resonance" between these two types of signals. In MFT, at every level in a hierarchy, the afferent signals are represented by the input signal field $\mathbf{X}$, and the efferent signals are represented by the modeling field signals $\mathbf{M}_h$; resonances correspond to high similarity measures $l(n|h)$ for some subsets of $\{n\}$ that are "recognized" as concepts (or objects) h. The mechanism leading to the resonances, Eqs. (3)–(6), is an elementary thought-process. In this process, subsets of signals corresponding to objects or situations are understood as concepts, signals acquire meanings.

Kant's three volumes on theory of the mind, Critique of Pure Reason, Critique of Judgment, and Critique of Practical Reason [40,106,107] describe the structure of the mind similarly to MFT. Pure reason or the faculty of understanding contains concept-models. The faculty of judgment, or emotions, establishes correspondences between models and data about the world acquired by sensory organs (in Kant's terminology, between general concepts and individual events). Practical reason contains models of behavior. Kant was first to recognize that emotions are an inseparable part of cognition. The only missing link in Kantian theory is the knowledge instinct. Kant underappreciated a pervading need for concept adaptation; he considered concepts as given a priori.

A dynamic aspect of working of the mind, described by dynamic logic, was first given by Aristotle [108]. He described thinking as a learning process in which an a priori form-as-potentiality (fuzzy model) meets matter (sensory signals) and becomes a form-as-actuality (a concept of the mind). He pointed out an important aspect of dynamic logic, reduction of fuzziness during learning: Forms-potentialities are fuzzy (do not obey logic), whereas forms-actualities are logical.

History preserved for us evidence of Aristotle's foresight. When Alexander the Great, Aristotelian pupil, was fighting in Persia, he wrote to his teacher: "Aristotle, I heard you are writing books now. Are you going to make our secret knowledge public?" In a reply letter Aristotle wrote: "Alexander, do not worry: nobody will understand" [109].

7.3. Understanding

In the elementary thought process, subsets in the incoming signals are associated with recognized models, which are *understood* as objects or situations. In other words *signal subsets* acquire *meaning*, for example, a subset of retinal signals acquires a meaning of a chair. There are several aspects to understanding and meaning. First, object-models are connected (by emotional signals [8,10]) to instincts that they might satisfy, and also to behavioral models that can make use of them for instinct satisfaction. Second, an object is understood in the context of a more general situation in the next hierarchical level consisting of more general concept-models, which accepts as input-signals the results of object recognition. That is, each recognized object-model sends (in neural terminology, activates) an output signal; and a set of these signals comprises input signals for the next level models, which "cognize" more general concept-models, like relations and situations. And this process continues up and up the hierarchy of the mind toward the most general models a system could come up with, such as models of universe (scientific theories), models of self (psychological concepts), models of meaning of existence (philosophical concepts), models of a priori transcendent intelligent subject (theological concepts).

7.4. Conscious and unconscious

Why is there consciousness? Why would a feature like consciousness appear in the process of evolution? The answer to this question seems clear: consciousness directs the will and results in a better adaptation for survival. In simple situations, when only minimal adaptation is required, an instinct directly wired to action is sufficient, and unconscious processes can efficiently allocate resources and will. However, in complex situations, when adaptation is complicated, various instincts might contradict one another. Undifferentiated unconscious psychic functions result in ambivalence and ambitendency; every position entails its own negation, leading to an inhibition. This inhibition cannot be resolved by unconscious that does not differentiate among alternatives. Direction is impossible without differentiation. Consciousness is needed to resolve an instinctual impasse by suppressing some processes and allocating power to others. By differentiating alternatives, consciousness can direct a psychological function to a goal.

We are not conscious about most of organism functioning; blood flow, breathing, workings of heart and stomach are unconscious, at least as long as they work as appropriate. The same is true about most of the processes in the brain and mind. We are not conscious about fuzzy models competing for evidence in retinal signals, etc. We become conscious about concepts, only during resonance, when a model-concept matches bottom-up signals and become crisp.

To put it more accurately, crisper models are better accessible by consciousness. In psychic functioning, evolutionary directions and our personal goals are to increase consciousness. But, this is largely unconscious, because our direct knowledge of ourselves is limited to consciousness. This fact creates a lot of confusion about consciousness. So, what is consciousness?

Consciousness is an awareness or perception of inward psychological facts, a subjective experience of sensing, feelings, or thoughts. This definition is taken from the Webster's Dictionary. But, a more detailed, scientific analysis of consciousness has proven to be difficult. For a long time it seemed obvious that consciousness completely pervades our entire mental life, or at least its main aspects. Now, we know that this idea is wrong, and the main reason for this misconception has been analyzed and understood: We are conscious only about what we are conscious of, and it is extremely difficult to notice anything else.

Misconceptions about consciousness noted by Jaynes [110] include: consciousness is nothing but a property of matter, or a property of living things, or a property of neural systems. These three "explanations" attempted to dismiss consciousness as an epiphenomenon, an unimportant quality of something else. They are useless because the problem is in *explaining* the relationships of consciousness to matter, to life, and to neural systems. These dismissals of consciousness are not very different from saying that there is no consciousness; but, of course, this statement refutes itself (if somebody makes such a statement unconsciously, there is no point of discussing it). A dualistic position is that consciousness belongs to the world of ideas and has nothing to do with the world of matter. But the scientific problem *is* in explaining the consciousness as a natural-science phenomenon; that is to relate consciousness and the material world. Searle [111] suggested that any explanation of consciousness has to account for it being real and based on physical mechanisms in the brain. Among properties of consciousness requiring explanation he listed unity and intentionality (we perceive our consciousness as being *unified* in the space of our perceptions and in the time of our life; consciousness is about something; this "about" points to its *intentionality*).

In his book on consciousness, Searle [112] reviewed recent attempts to explain consciousness, and came to the conclusion that little progress was made during the 1990s. Penrose [5] suggested that consciousness cannot be explained by known physical laws of matter. His arguments descend from Gödel's proofs of inconsistency and incompleteness of logic. We have already mentioned that this, however, only proves [28] that the mind is not a system of logical rules.

Knowledge of consciousness is primarily of introspective origin. Understanding of consciousness requires differentiating conscious and unconscious psychic processes, so we need to understand what is psychic, what is unconscious, and what is consciousness. Our experiences can be divided into somatic and psychic. A will modifying instinctual reflexes indicates a presence of psyche, but not necessarily consciousness. Often, we associate consciousness with a subjective perception of free will. Consciousness about somatic experiences is limited by the unknown in the outer world. Similarly, consciousness about psychic experiences is limited by *the unknown in psyche, or unconscious*. Roughly speaking, there are three conscious/unconscious levels of psychic contents:

(1) contents that can be recalled and made conscious voluntarily (memories);
(2) contents that are not under voluntary control, we know about them because they spontaneously irrupt into consciousness; and
(3) contents inaccessible to consciousness.

We know about the latter through scientific deductions.

Consciousness is not a simple phenomenon, but a complicated differentiated process. Jung differentiated four types of consciousness related to experiences of feelings (emotions), thoughts (concepts), sensations, and intuitions [42]. In addition to these four psychic functions, consciousness is characterized by the attitude: Introverted, concentrated mainly on the inner experience, or extroverted, concentrated mainly on the outer experience. Interplay of various conscious and unconscious levels of psychic functions and attitudes results in a number of types of consciousness; interactions of these types with individual memories and experiences make consciousness dependent on the entire individual experience producing variability among individuals. The reviewed theories of the mind only touched on relationships between concepts and consciousness. An idea that better differentiated, crisper model-concepts are more conscious is close to Jung's views. Mechanisms of other types of consciousness are less understood and their mathematical descriptions belong to future. Future research would also address emergence in evolution of different types of consciousness, elaborating on Jungian ideas.

Totality and undividedness of consciousness are important adaptive properties needed to concentrate power on the most important goal at every moment. This is illustrated, for example, by clinical cases of divided consciousness and multiple personalities, resulting in maladaptation up to a complete loss of functionality. Simple consciousness needs only to operate with relatively few concepts. Humans need more differentiation for selecting more specific goals in more complex environment. The scientific quest is to explain these opposite tendencies of consciousness: how does consciousness pursues undividedness and differentiation at once? There is no mystery, the knowledge instinct together with the hierarchical structure of the mind hold the key to the answer. Whereas every level pursues differentiation, totality belongs to the highest levels of the hierarchy. Future research will have to address these mechanisms in their fascinating details.

Intentionality is a property of referring to something else, and consciousness is about something. This "aboutness" many philosophers refer to as intentionality. In everyday life, when we hear an opinion we do not just collate it in our memory and relate to other opinions (like a pseudo-scientist in a comedy); this would not lead very far. We wish to know what are the aims and intentions associated with this opinion. Mechanisms of perceiving intent versus specific words were studied by Valerie Reyna and Charles Brainerd, who discuss the contrast between *gist* and *verbatim* systems of memory and decision making [113]. Often, we perceive the *intent* of what are said better then specific words, even if the words are chosen to disguise the intent behind causal reasoning. The desire to know and the ability to perceive the goal indicates that in psyche, *final standpoint or purpose* is more important than the *causal* one. This intentionality of psyche was already emphasized by Aristotle in his discussions of the end cause of forms of the mind [108]. Intentionality of consciousness is more fundamental than "aboutness", it is *purposiveness* [114].

The intentional property of consciousness led many philosophers during the last decades to believe that intentionality is a unique and most important characteristic of consciousness: according to Searle, only conscious beings could be intentional. But, the mechanism of the knowledge instinct leads to an opposite conclusion. Intentionality is a fundamental property of life: even a simplest living being is a result of long evolution and its every component, say a gene, or a protein has a purpose and intent. In particular, every model-concept has evolved with an intent or purpose to recognize a particular type of signal (event, message, concept) and to act accordingly (e.g., send recognition message to other parts of the brain and to behavioral models). Aristotle was the first to explain the intentionality of the mind this way; he argued that intentionality should be explained through the a priori contents of the mind [115]. Possibly, future theoretical developments of mechanisms of the knowledge instinct will explain the minds intentionality and purposiveness in its complexity.

Is there any specific relationship between consciousness and intentionality? If so, it is just the opposite of Searle's hypothesis of intentionality implying consciousness. Affective, subconscious, lower-bodily-level emotional responses are concerned with immediate survival, utilitarian goals, and therefore are *intentional in the most* straightforward way. A higher-intellectual-level consciousness is not concerned with the immediate survival, but with the overall understanding of the world, with knowledge and beauty; it can afford to be impartial, abstract, and less immediately-intentional than the rest of the psyche; its intentions might be directed toward meanings and purposes of life. As we discuss few pages below, the highest creative aspect of individual consciousness and the abilities of perceiving beautiful and sublime are intentional without any specific, lower-level utilitarian goal, they are intentional toward self-realization, toward future-self beyond current-self. Due to the current mathematical theories reviewed in this article we can more accurately manipulate these metaphorical descriptions to obtain solutions to long-standing philosophical problems. In addition, we can identify directions for concrete studies of these metaphors in future mathematical simulations and laboratory experiments.

Unity of consciousness refers to conscious mental states being parts of a unified sequence and simultaneous conscious events are perceived as unified into a coherent picture. Searle's unity is close to what Kant called "the transcendental unity of apperception". In MFT, this internal perception is explained as all perceptions, due to a property of the special model involved in consciousness, called Ego by psychologists. The properties of Ego-model explain the properties of consciousness. When certain properties of consciousness seem difficult to explain, we should follow the example of Kant, we should turn the question around and ask: Which properties of Ego model would explain the phenomenological properties of consciousness?

Let us begin the analysis of the structures of the Ego-model and the process of its adaptation to the constantly changing world, from evolutionary-preceding simpler forms. What is the initial state of consciousness: an undifferentiated unity or a "booming, buzzing confusion" [116]? Or, let us make a step back in the evolutionary development and ask, what is the initial state of pre-conscious psyche? Or, let us move back even further toward evolution of sensory systems and perception. When building a robot for a factory floor, why provide it with a sensor? Obviously, such an expensive thing as a sensor is needed to achieve specific goals: to sense the environment with the purpose to accomplish specific tasks. Providing a robot with a sensor goes together with an ability to utilize sensory data.

Similarly, in the process of evolution, sensory abilities emerged together with perception abilities. A natural evolution of sensory abilities could not result in a "booming, buzzing confusion", but must result in evolutionary advantageous abilities to avoid danger, attain food, etc. Primitive perception abilities (observed in primitive animals) are *limited* to few types of concept-objects (light-dark, warm-cold, edible-nonedible, dangerous-attractive ...) and are directly "wired" to proper actions. When perception functions evolve further, beyond immediate actions, it is through the development of complex internal model-concepts, which unify simpler object-models into a unified and flexible model of the world. Only at this point of possessing relatively complicated differentiated concept-models composed of a large number of sub-models, an intelligent system can experience a "booming, buzzing confusion", if it faces a new type of environment. A primitive system is simply incapable of perceiving confusion: It perceives only those "things" for which it has concept-models and if its perceptions do not correspond to reality, it just does not survive without experiencing confusion. When a baby is born, it undergoes a tremendous change of environment, most likely without much conscious confusion. The original state of consciousness is undifferentiated unity. It possesses a single modality of primordial undifferentiated Self-World.

The initial unity of psyche limited abilities of the mind, and further development proceeded through differentiation of psychic functions or modalities (concepts, emotions, behavior); they were further differentiated into multiple concept-models, etc. This accelerated adaptation. Differentiation of consciousness is a relatively recent process [110, 120].

Consciousness is about aspects of concept-models (of the environment, self, past, present, future plans, and alternatives) and emotions [117] to which we can direct our attention. As already mentioned, MFT explains consciousness as a specialized Ego-model. Within this model, consciousness can direct attention at will. This conscious control of will is called the free will. A subjective feeling of free will is a most cherished property of our psyche. Most of us feel that this is what makes us different from inanimate objects and simple forms of life. And this property is a most difficult one to explain rationally or to describe mathematically. But, let us see how far we can go towards understanding this phenomenon. We know that raw percepts are often not conscious. As mentioned already, for example, in the visual system, we are conscious about the final processing stage, the integrated crisp model, and unconscious about intermediate processing. We are unconscious about eye receptive fields; about details of visual perception of motion and color as far as it takes place in our brain separately from the main visual cortex, etc. [66]. In most cases, we are conscious only about the integrated scene, crisp objects, etc.

These properties of consciousness follow from properties of concept-models, they have conscious (crisp) and unconscious (fuzzy) parts, which are accessible and inaccessible to consciousness, that is to Ego-model. In pre-scientific literature about mechanisms of the mind there was a popular idea of homunculus, a little mind, inside our mind, which perceived our perceptions and made them available to our mind. This naive view is amazingly close to actual scientific explanation. The fundamental difference is that the scientific explanation does not need an infinite chain of homunculi inside homunculi. Instead, there is a hierarchy of the mind models with their conscious and unconscious aspects. The higher in the hierarchy, the less is the conscious differentiated aspect of the models. Until at the top of the hierarchy there are mostly unconscious models of the meaning of our existence (which we discuss later).

Our internal perceptions of consciousness due to Ego-model "perceive" crisp conscious parts of other models similar to models of perception "perceive" objects in the world. The properties of consciousness as we perceive them, such as continuity and identity of consciousness, are due to properties of the Ego-model. What is known about this "consciousness"-model? Since Freud, a certain complex of psychological functions was called Ego. Jung considered Ego to be based on a more general model or archetype of Self. Jungian archetypes are psychic structures (models) of a primordial origin, which are mostly inaccessible to consciousness, but determine the structure of our psyche. In this way, archetypes are similar to other models, e.g., receptive fields of the retina are not consciously perceived, but determine the structure of visual perception. The Self archetype determines our phenomenological subjective perception of ourselves, and in addition, structures our psyche in many different ways, which are far from being completely understood. An important phenomenological property of Self is the perception of uniqueness and indivisibility (hence, the word *individual*).

Consciousness, to a significant extent, coincides with the conscious part of the archetype-model of Self. A conscious part of Self belongs to Ego. Not everything within Ego (as defined by Freud) is conscious. Individuality as a total character distinguishing an individual from others is a main characteristic of Ego. Not all aspects of individuality are conscious, so, the relationships among the discussed models can be summarized to some extent, as:

$$\text{Consciousness} \in \text{Individuality} \in \text{Ego} \in \text{Self} \in \text{Psyche}.$$

The sign "$\in$" here means "is a part of". Consciousness-model is a subject of free will; it possesses, controls, and directs free will. Free will is limited by laws of nature in the outer world and in the inner world by the unconscious aspects of Self. Free will belongs to consciousness, but not to the conscious and unconscious totality of the psyche. Clearly, much of the above discussion still has a long way to go to mathematical modeling and laboratory experimentation; but it does not seem anymore as mystery beyond future physics of the mind.

Many contemporary philosophers consider subjective nature of consciousness to be an impenetrable barrier to scientific investigation. Chalmers differentiated hard and easy questions about consciousness [118] as follows. Easy questions, that will be answered better and better, are concerned with brain mechanisms: which brain structures are responsible for consciousness? Hard questions, that no progress can be expected about, are concerned with the subjective nature of consciousness and *qualia*, subjective feelings associated with every conscious perception. Nagel described it dramatically with a question: "What is it like to be a bat?" [119] But I disagree. I do not think these questions are hard. These questions are not mysteries; they are just wrong questions for a scientific theory. Newton, while describing the laws of planet motion, did not ask: "What is it like to be a planet?" (even so, something like this feeling *is* a part of scientific intuition). The subjective nature of consciousness is not a mystery. It is explained due to the subjective nature of the concept-models that we are conscious of. The subjectivity is the result of combined a priority and adaptivity of the consciousness-model, the unique genetic a priori structures of psyche together with our unique individual experiences. I consider the only hard questions about consciousness to be *free will and the nature of creativity*.

Let us summarize. Most of the mind's operations are not accessible to consciousness. We definitely know that neural firings and connections cannot be perceived consciously. In the foundations of the mind there are material processes in the brain inaccessible to consciousness. Jung suggested that conscious concepts are developed by the mind based on genetically inherited structures, archetypes, which are inaccessible to consciousness [42,120]. Grossberg [9] suggested that only signals and models attaining a resonant state (that is signals matching models) can reach consciousness. It was further detailed by Taylor [121]; he related consciousness to the mind being a control mechanism of the mind and body. A part of this mechanism is a prediction model. When this model predictions differ from sensory observations, this difference may reach a resonant state, which we are consciousness about. To summarize the above analyses, the mind mechanisms, described in MFT by dynamic logic and fuzzy models, are not accessible to consciousness. Final results of dynamic logic processes, resonant states characterized by crisp models and corresponding signals are accessible to consciousness.

7.5. Imagination

Imagination involves excitation of a neural pattern in a sensory cortex in absence of an actual sensory stimulation. For example, visual imagination involves excitation of visual cortex, say, with closed eyes [9,66]. Imagination was long considered a part of thinking processes; Kant [40] emphasized the role of imagination in the thought process, he called thinking "a play of cognitive functions of imagination and understanding". Whereas pattern recognition and artificial intelligence algorithms of recent past would not know how to relate to this [3,4], Carpenter and Grossberg resonance model [57] and the MFT dynamics both describe imagination as an inseparable part of thinking. Imagined patterns are top–down signals that *prime* the perception cortex areas (*priming* is a neural terminology for making neurons to be more readily excited). In MFT, the imagined neural patterns are given by models $\mathbf{M}_h$.

Visual imagination, as mentioned, can be "internally perceived" with closed eyes. The same process can be mathematically modeled at higher cognitive levels, where it involves models of complex situations or plans. Similarly, models of behavior at higher levels of the hierarchy can be activated without actually propagating their output signals down to actual muscle movements and to actual acts in the world. In other words, behavior can be imagined, along with its consequences, it can be evaluated, and this is the essence of plans. Sometimes, imagination involves detailed alternative courses of actions considered and evaluated consciously. Sometimes, imagination may involve fuzzy or vague, barely conscious models, which reach consciousness only after they converge to a "reasonable" course of action, which can be consciously evaluated. From a mathematical standpoint, this latter mechanism must predominate, conscious evaluation cannot involve all possible courses of action; it would lead to combinatorial complexity and impasse. It remains to be proven in brain studies, which will identify the exact brain regions and neural mechanisms involved.

In agreement with neural data, MFT adds details to Kantian description: thinking is a play of top–down *higher-hierarchical-level* imagination and bottom–up *lower-level* understanding. Kant identified this "play" [described by (3)–(6)] as a source of aesthetic emotion. Kant used the word "play", when he was uncertain about the exact mechanism; this mechanism, according to our suggestion, is the knowledge instinct and dynamic logic.

7.6. Bodily instincts and emotions

The functioning of the mind and brain cannot be understood in isolation from the system's "bodily needs". For example, a biological system (and any autonomous system) needs to replenish its energy resources (eat). This and other fundamental unconditional needs are indicated to the system by instincts and emotions. As we discussed, scientific terminology in this area is still evolving; for our purpose of making a step toward uncovering neural mechanisms of the mind, we describe instincts mathematically as internal sensors, which measurements directly indicate unconditional needs of an organism. For example, instinct for food measures sugar level in blood, and related emotional signals are perceived by psyche as "hunger"; they activate behavioral models related to food searching and eating. Bodily instinctual influences on workings of the mind modify the object-perception process (3)–(6) in such a way that desired objects get enhanced recognition. This is the reason a hungry person "sees food all around". In MFT it can be accomplished by modifying priors, $r(h)$ in Eqs. (2), (3) according to the degree to which an object of type h can satisfy a particular instinct [122]. Details of these mechanisms are not considered here.

7.7. Aesthetic emotions and the instinct for knowledge

Recognizing objects in the environment and understanding their meaning is so important for survival that a special instinct evolved for this purpose. This instinct for learning and improving concept-models I call the instinct for knowledge. In MFT it is described by maximization of similarity between the models and the world, Eq. (2). Emotions related to satisfaction–dissatisfaction of this instinct are perceived by us as harmony–disharmony (between our understanding of how things ought to be and how they actually are in the surrounding world). According to Kant [40] these are aesthetic emotions (emotions that are not related directly to satisfaction or dissatisfaction of bodily needs).

The instinct for knowledge makes little kids, cubs, and piglets jump around and play fight. Their inborn models of behavior must adapt to their body weights, objects, and animals around them long before the instincts of hunger and fear will use the models for direct aims of survival. Childish behavior just makes the work of the knowledge instinct more observable; to varying degrees, this instinct continues acting throughout our lives. All the time we are bringing our internal models into correspondence with the world. In adult life, when our perception and understanding of the surrounding world is adequate, aesthetic emotions are barely perceptible: the mind just does its job. Similarly, we do not usually notice adequate performance of our breathing muscles and satisfaction of the breathing instinct. However, if breathing is difficult, negative emotions immediately reach consciousness. The same is true about the knowledge instinct and aesthetic emotions: if we do not understand the surroundings, if objects around do not correspond to our expectations, negative emotions immediately reach consciousness. We perceive these emotions as disharmony between our knowledge and the world. Thriller movies exploit the instinct for knowledge: they are mainly based on violating our expectations; their personages are shown in situations, when knowledge of the world is inadequate for survival.

Let me emphasize again, aesthetic emotions are not peculiar to art and artists, they are inseparable from every act of perception and cognition. In everyday life we usually do not notice them. Aesthetic emotions become noticeable at higher cognitive levels in the mind hierarchy, when cognition is not automatic, but requires conscious effort. Damasio view [46] of emotions defined by visceral mechanisms, as far as discussing higher cognitive functions, seems erroneous in taking secondary effects for the primary mechanisms. People often devote their spare time to increasing their knowledge, even if it is not related to their job and a possibility of promotion. Pragmatic interests could be involved: knowledge makes us more attractive to friends and could help find sexual partners. Still, there is a remainder, a pure joy of knowledge, aesthetic emotions satisfying the knowledge instinct.

7.8. Beautiful and sublime

Contemporary cognitive science is at a complete loss when trying to explain the highest human abilities, the most important and cherished abilities to create and perceive beautiful and sublime experiences. Their role in the working of the mind is not understood. MFT explains that simple harmony is an elementary aesthetic emotion related to improvement of object-models. Higher aesthetic emotions are related to the development and improvement of more complex "higher" models at higher levels of the mind hierarchy. The highest forms of aesthetic emotion are related to the most general and most important models near the top of the hierarchy. According to Kantian analysis [40,123], among the highest models are models of the meaning of our existence, of our purposiveness or intentionality, and beauty is related to improving these models.

Models of our purposiveness are largely fuzzy and unconscious. Some people, at some points in their life, may believe that their life purpose is finite and concrete, for example to make a lot of money, or build a loving family and bring up good children. These models are aimed at satisfying powerful instincts, but not the knowledge instinct, and they do not reflect the highest human aspirations. Everyone who has achieved a finite goal of making money or raising good children knows that this is not the end of his or her aspirations. The reason is that everyone has an ineffable feeling of partaking in the infinite, while at the same time knowing that our material existence is finite. This contradiction cannot be resolved. For this reason models of our purpose and meaning cannot be made crisp and conscious, they will forever remain fuzzy and partly unconscious.

Everyday life gives us little evidence to develop models of meaning and purposiveness of our existence. People are dying every day and often from random causes. Nevertheless, life itself demands belief in one's purpose; without such a belief it is easier to get drunk or take drugs than to read this article. These issues are not new; philosophers and theologists expounded them from time immemorial. The knowledge instinct theory gives us a scientific approach to the eternal quest for the meaning. We perceive an object or a situation as beautiful, when it stimulates improvement of the highest models of meaning. Beautiful is what "reminds" us of our purposiveness. This is true about perception of beauty in a flower or in an art object. Just an example, R. Buckminster Fuller, an architect, best known for inventing the geodesic dome wrote: "When I'm working on a problem, I never think about beauty. I think only how to solve the problem. But when I have finished, if the solution is not beautiful, I know it is wrong." [124]. The MFT explanation of the nature of beautiful helps understanding an exact meaning of this statement and resolves a number of mysteries and contradictions in contemporary aesthetics [125,126].

The feeling of spiritually sublime is similar and different from beautiful. Whereas beautiful is related to improvement of the models of *cognition*, sublime is related to improvement of the models of *behavior* realizing the highest meaning in our life. Beautiful and sublime are not finite. MFT tells us that mathematically, improvement of complex models is related to choices from infinite number of possibilities. A mathematician may consider 100^{100}, or million to the millionth power as a finite number. But for a physicist, a number that exceeds all elementary events in the life of the Universe is infinite. A choice from infinity is infinitely complex and contains infinite information. Therefore, choices of beautiful and sublime contain infinite information. This is not a metaphor, but exact mathematical fact. Beauty is at once objective and subjective. It really exists, cultures and individuals cannot exist without the ability for beauty, and still, it cannot be described by any finite algorithm or a set of rules.

Beauty of a physical theory discussed sometimes by physicists is similar in its infinity to beauty in an artwork. For a physicist, beauty of a physical theory is related to improving the models of the meaning in our understanding of the universe. This satisfies a scientist's quest for the purpose, which he identifies with the purpose in the world.

7.9. Intuition

Intuitions include inner perceptions of object-models, imaginations produced by them, and their relationships with objects in the world. They include also higher-level models of relationships among simpler models. Intuitions involve fuzzy unconscious concept-models, which are in a state of being formed, learned, and being adapted toward crisp and conscious models (say, a theory). Conceptual contents of fuzzy models are undifferentiated and partly unconscious. Similarly, conceptual and emotional contents of these fuzzy mind states are undifferentiated; concepts and emotions are mixed up. Fuzzy mind states may satisfy or dissatisfy the knowledge instinct in varying degrees before they become differentiated and accessible to consciousness, hence the vague complex emotional-cognitive feel of an intuition. Contents of intuitive states differ among people, but the main mechanism of intuition is the same among artists and scientists. Composer's intuitions are mostly about sounds and their relationships to psyche. Painter's intuitions are mostly about colors and shapes and their relationships to psyche. Writer's intuitions are about words, or more generally, about language and its relationships to psyche. Mathematical intuition is about structure and consistency within a theory, and about relationships between the theory and a priori content of psyche. Physical intuition is about the real world, first principles of its organization, and mathematics describing it.

7.10. Language, cognition, and symbols

Why is the word "symbol" used in such opposite ways: to denote trivial objects, like traffic signs or mathematical notations, and also to denote objects affecting entire cultures over millennia, like Magen David, Cross, or Crescent?

Let us compare in this regard opinions of two founders of contemporary semiotics, Charles Peirce and Ferdinand De Saussure. Peirce classified signs into symbols, indexes, and icons [127]. Icons have meanings due to resemblance to the signified (objects, situations, etc.), indexes have meanings by direct connection to the signified, and *symbols* have meaning due to arbitrary conventional agreements. Saussure used different terminology, he emphasized that the *sign* receives meaning due to arbitrary conventions [128], whereas *symbol* implies motivation.

Both Peirce and Saussure wanted to understand the process in which signs acquire meanings. Both of them failed: workings of the mind were not known at the time. Consider Peircian icons; they resemble objects or situations because of specific mechanisms of perception and recognition in our mind. These mechanisms should be analyzed and understood as an essential part of meaning creation. Peircian assumption that icons in themselves resemble situations in the world is too simplistic. Algorithms based on this assumption led to irresolvable difficulties related to combinatorial complexity. Similarly, arbitrariness emphasized by Peirce and Saussure did not lead to algorithms of meaning creation. Arbitrary signs have no grounding in real world. Meanings cannot be created by unmotivated choices on the interconnections of arbitrary signs, this type of choices lead to combinatorial complexity. In infinite systems, they lead to Gödelian contradictions. Similarly, mechanisms of meaning creation were not found by founders of "symbolic artificial intelligence", when they used the motivationally loaded word "symbol" for arbitrary mathematical notations. Mathematical notations, just because they are called symbols, do not hold a key to the mystery of cultural and psychological symbols. Multiple meanings of the word "symbol" misguided their intuition. This is an example of what Wittgenstein called "bewitchment by language".

The MF theory emphasizes that meaning creation consists in bringing unconscious into consciousness in the process of model adaptation. This process is "motivated" by the instinct for knowledge. The motivated meaning creation, connecting conscious and unconscious, is consistent with Jungian explanations of the nature of symbols [120]. This motivates to use the word *symbol* for the processes of meaning creation, and to use the word *sign* for conventional or nonadaptive entities. This corresponds to Pribram's [129] interpretation of signs, as nonadaptive neural signals with fixed meanings.

Meanings are created by symbol-processes in the mind. Language plays special role in these processes. Language accumulates cultural knowledge of the world. Through communication among people, language provides grounding for abstract model-concepts at higher levels in the mind hierarchy. The mechanism of this relationship between language and cognition is joint language-cognitive models. These joint models are organized in parallel hierarchies of language models (words, texts) and cognitive models (world representations in the mind). Near the bottom of these hierarchies words refer to objects. Higher up, complex texts refer to complex situations. An amazing result of the described mechanism is that words within texts refer to objects within situations, and this reference at higher levels corresponds to the words-objects relationships at lower levels. Because of this multi-level hierarchical structure, maintaining meaningful relationships throughout the hierarchy, language is a coherent structure and not a set of arbitrary notations for arbitrary relationships. This meaning–maintaining hierarchy makes possible "the infinite use of finite means". We do not know to which extent the hierarchies are inborn or created by mechanisms, which construct higher levels from lower ones. Brighton et al. [102] results could be interpreted in the following way: a higher level is predicted from a lower one. Future research will study this conjecture in more details [105,130].

Cultural evolution results in selection and preservation in language of important meanings. Importance of meanings of various models and texts is based in culture; meanings and their importance are "biased" by culture; these biases constitute culture. But the deconstruction idea that meanings are arbitrary is unscientific. Scientific quest is to explain creation of meanings and the reviewed research makes steps in this direction.

A symbol-process involves conscious and unconscious, concepts and emotions, inborn models-archetypes and models learned from culture, language, and cognition. Symbol processes continue up and up the hierarchy of models and mind toward the most general models. Due to language, they continue in culture through many generations. In semiotics this process is called *semiosis*, a continuous process of creating and interpreting the world outside (and inside our mind). Symbols are processes creating meanings.

7.11. Creativity, differentiation, and synthesis

Creativity is an ability to improve and create new model-concepts. In a small degree it is present in everyday perception and cognition. Usually the words "creativity", "creative", or "discovery" are applied to improving or creating new model-concepts at higher cognitive levels, concepts that are important for the entire society or culture. A crisp and specific model could only match a specific content; therefore it cannot lead to creation of new contents. Creativity and discovery, according to Section 5, involve vague, fuzzy models, which are made more crisp and clear. It occurs, therefore, at the border between consciousness and unconscious. A similar nature of creative process, involving consciousness and unconscious, was discussed by Jung [120]. Creativity usually involves intuition, as discussed above: fuzzy undifferentiated feelings-concepts.

Creativity is driven by the knowledge instinct. Two main mechanisms of creativity, the components of the knowledge instinct, are differentiation and synthesis. Differentiation is a process of creating new, more specific and more detailed concept-models from simpler, less differentiated and less conscious models. Mathematical mechanisms of differentiation were discussed in Section 5. The role of language in differentiation of cognition was discussed in Section 6, as mentioned, this research is in its infancy and a subject of future research.

Synthesis is a process of connecting detailed crisp concept-models to the unconscious, instincts, and emotions. The need for synthesis comes from the fact that most of our concept-models are acquired from language. The entire conceptual content of the culture is transmitted from generation to generation through language; cognitive concept-models cannot be transmitted directly from brain to brain. Therefore, concepts acquired from language have to be used by individual minds to create cognitive concepts. The mechanism of integrating cognition and language, discussed in Section 6, explains that language concepts could be detailed and conscious; but not necessarily connected to equally detailed cognitive concepts, to emotions, and to the knowledge instinct. Connecting language and cognition involves differentiating cognitive models, developing cognitive models, whose differentiation and consciousness approaches that of language models. Every child acquires language between one and seven, but it takes the rest of life to connect abstract language models to cognitive concept-models, to emotions, instincts, and to the life's needs. This is the process of synthesis; it integrates language and cognition, concepts and emotions, conscious and unconscious, instinctual and learned. Current research directions discussed in Section 6 are just touching on these mechanisms of synthesis. It is largely an area for future research.

Another aspect of synthesis, essential for creativity, is developing a unified whole within psyche, a feel and intuition of purpose and meaning of existence. It is necessary for concentrating will, for survival, for achieving individual goals, and in particular for satisfying the knowledge instinct by differentiating knowledge. Concept-models of purpose and meaning, as discussed are near the top of the mind hierarchy; they are mostly unconscious and related to feelings of beautiful and sublime. A condition of synthesis is correspondence among a large number of concept-models. A knowledge instinct as discussed in Section 5 is a single measure of correspondence between all the concept-models and all the experiences-data about the world. This is, of course, a simplification. Certain concept-models have high value for psyche (e.g., family, success, certain political causes) and they affect recognition and understanding of other concepts. This is a mechanism of differentiation of the knowledge instinct. Satisfaction of the knowledge instinct therefore is not measured by a single aesthetic emotion, but by a large number of aesthetic emotions. The entire wealth of our knowledge should be brought into correspondence with itself, this requires a manifold of aesthetic emotions. Differentiation of emotions is performed by music [126], but this is beyond the scope of the review.

There is an opposition between differentiation and synthesis in individual minds as well as in the collective psyche. This opposition leads to complex evolution of cultures. Differentiated concepts acquire meaning in connections with instinctual and unconscious, in synthesis. In evolution of the mind, differentiation is the essence of the development of the mind and consciousness, but it may bring about a split between conscious and unconscious, between emotional and conceptual, between language and cognition. Differentiated and refined models existing in language may loose connection with cognitive models, with people's instinctual needs. If the split affects collective psyche, it leads to a loss of the creative potential of a community or nation. This was the mechanism of death of great ancient civilizations. The development of culture, the very interest of life requires *combining differentiation and synthesis*. Evolution of the mind and cultures is determined by this complex nonlinear interaction: One factor prevails, then another [126]. This is an area for future research.

7.12. Teleology, causality, and the knowledge instinct

Teleology explains the Universe in terms of purposes. In many religious teachings, it is a basic argument for the existence of God: If there is purpose, an ultimate Designer must exist. Therefore, teleology is a hot point of debates between creationists and evolutionists: Is there a purpose in the world? Evolutionists assume that the only explanation is causal. Newton laws gave a perfect causal explanation for the motion of planets: A planet moves from moment to moment under the influence of a gravitational force. Similarly, today science explains motions of all particles and fields according to causal laws, and there are exact mathematical expressions for fields, forces and their motions. Causality explains what happens in the next moment as a result of forces acting in the previous moment. Scientists accept this causal explanation and oppose to teleological explanations in terms of purposes. The very basis of science, it seems, is on the side of causality, and religion is on the side of teleology.

However, at the level of the first physical principles this is wrong. The contradiction between causality and teleology does not exist at the very basic level of fundamental physics. The laws of physics, from classical Newtonian laws to quantum superstrings, can be formulated equally as causal or as teleological. An example of teleological principle in physics is energy minimization, particles move so that energy is minimized. As if particles in each moment know their purpose: to minimize the energy. The most general physical laws are formulated as minimization of action. Action is a more general physical entity than energy; it is an intuitive name for a mathematical expression called Lagrangian. Causal dynamics, motions of particles, quantum strings, and superstrings are determined by minimizing Lagrangian-action [131]. A particle under force moves from point to point as if it knows its final purpose, to minimize Lagrangian-action. Causal dynamics and teleology are two sides of the same coin.

The knowledge instinct is similar to these most general physical laws: evolution of the mind is guided by maximization of knowledge. A mathematical structure of similarity (2) or its continuous version [59] is similar to Lagrangian, and it plays a similar role; it bridges causal dynamic logic of cognition and teleological principle of maximum knowledge. Similarly to fundamental physics, dynamics and teleology are equivalent: Dynamic logic follows from maximization of knowledge and vice versa. Ideas, concept-models change under the “force” of dynamic logic, as if they know the purpose: Maximum knowledge. One does not have to choose between scientific explanation and teleological purpose: Causal dynamics and teleology are equivalent.

7.13. Mind and brain, experimental evidence

Historically, the mind is described in psychological and philosophical terms, whereas the brain is described in terms of neurobiology and medicine. Within scientific exploration the mind and brain are different description levels of the same system. Establishing relationships between these descriptions is of great scientific interest. Today we approach solutions to this challenge [132], which eluded Newton in his attempt to establish physics of “spiritual substance” [133]. Detailed discussion of established relationships between the mind and brain is beyond the scope of this review. We briefly mention the main known and unknown facts and give references for future reading. Adaptive modeling abilities are well studied with adaptive parameters identified with synaptic connections [134]; instinctual learning mechanisms have been studied in psychology and linguistics [68,76,98,135]. General neural mechanisms of the elementary thought process (which are similar in MFT and ART [57]) include neural mechanisms for bottom–up (sensory) signals, top–down imagination model-signals, and the resonant matching between the two; these have been confirmed by neural and psychological experiments [136]. Ongoing research address relationships between neural processes and consciousness [121,132,137]. Relating MFT to brain mechanisms in details is a subject of ongoing and future research.

7.14. Predictions and testing

Ongoing and future research will confirm, disprove, or suggest modifications to specific mechanisms considered in Sections 5 and 6. These mechanisms include model parameterization and parameter adaptation, reduction of fuzziness during learning, and the similarity measure described by Eq. (2) as a foundation of the knowledge instinct and aesthetic emotion. Other mechanisms include on one hand, relationships between psychological and neural mechanisms of learning and, on the other hand, aesthetic feelings of harmony and emotions of beautiful and sublime. Future research will also investigate the validity of the dual integrated structure of model-concepts described by Eq. (9) as a foundation for interaction between cognition and language and for symbolic ability. A step in this direction will be to demonstrate in simulations that this mechanism actually integrates cognition and language without combinatorial complexity. Specific neural systems will need to be related to mathematical descriptions as well as to psychological descriptions in terms of subjective experiences and observable behavior. Ongoing simulation research addresses the evolution of models jointly with the evolution of language [138]. Also being investigated are the ways that MFT and the knowledge instinct relate to behavioral psychology and to the specific brain areas involved in emotional reward and punishment during learning [139]. Interesting unsolved problems include: detailed mechanisms of interactions between cognitive hierarchy and language hierarchy [60,61]; differentiated forms of the knowledge instinct, the infinite variety of aesthetic emotions perceived in music, their relationships to mechanisms of synthesis [126]; and interactions of differentiation and synthesis in the development of the mind during cultural evolution. Future experimental research will need to examine, in detail, the nature of hierarchical interactions, including mechanisms of learning hierarchy, to what extent the hierarchy is inborn vs. adaptively learned, and the hierarchy of the knowledge instinct.

Acknowledgements

It is my pleasure to thank people whose thoughts, ideas, encouragement, and support shaped this review: R. Brockett, G. Carpenter, A. Chernyakov, R. Deming, V. Dmitriev, F. Fontanari, M. Frank-Kamenetskii, W. Freeman, K. Fukunaga, L. Garvin, I. Ginzburg, R. Gudwin, S. Greineder, S. Grossberg, M. Karpovsky, S. Kashin, D. Levine, L. Levitin, R. Linnehan, T. Luginbuhl, A. Meystel, S. Middleton, K. Moore, C. Mutz, A. Ovsich, R. Payne, N. Pechersky, I. Perlovsky, V. Petrov, C. Plum, J. Rumer, E. Schakhnovich, W. Schoendorf, N. Shokhirev, J. Sjogren, D. Skatrud, R. Streit, E. Taborsky, I. Ternovsky, T. Ternovsky, E. Tichovolsky, B. Weijers, D. Vinkovetsky, Y. Vinkovetsky, P. Werbos, M. Xiarhos, L. Zadeh, G. Zainiev. I am thankful to anonymous referees for many suggestions that benefited this review.

References

[1] For the discussions and further references see: Grossberg S. Neural networks and natural intelligence. Cambridge, MA: MIT Press; 1988;
Albus JS, Meystel AM. Engineering of mind: An introduction to the science of intelligent systems. New York: Wiley; 2001;
Meystel AM, Albus JS. Intelligent systems: Architecture, design, and control. New York: Wiley; 2001;
Perlovsky LI. Neural networks and intellect: Using model based concepts. New York: Oxford University Press; 2001.
[2] The American Heritage College Dictionary. 3rd ed. Boston, MA: Houghton Mifflin; 2000. I would emphasize, these general dictionary formulations could only be taken as a starting point.
[3] Newell A. Intellectual issues in the history of artificial intelligence. In: Machlup F, Mansfield U, editors. The study of information. New York: J. Wiley; 1983.
[4] Minsky M. The society of mind. Cambridge, MA: MIT Press; 1988.
[5] Penrose R. Shadows of the mind. Oxford: Oxford University Press; 1994.
[6] Josephson B. An integrated theory of nervous system functioning embracing nativism and constructivism. In: International complex systems conference. Nashua, NH.; September 21–26, 1997.
[7] Hameroff S. Toward a scientific basis for consciousness. Cambridge, MA: MIT Press; 1994.
[8] Perlovsky LI. Neural networks and intellect: Using model based concepts. New York: Oxford University Press; 2001.
[9] Grossberg S. Neural networks and natural intelligence. Cambridge, MA: MIT Press; 1988.
[10] Grossberg S, Levine DS. Neural dynamics of attentionally modulated Pavlovian conditioning: Blocking, inter-stimulus interval, and secondary reinforcement. Psychobiology 1987;15(3):195–240.
[11] Zadeh LA. Information granulation and its centrality in human and machine intelligence. In: Proceedings of the conf on intelligent systems and semiotics '97. Gaithersburg, MD; 1997. p. 26–30.
[12] Meystel A. Semiotic modeling and situational analysis. Bala Cynwyd, PA: AdRem; 1995.
[13] Edelman GM, Tononi G. A universe of consciousness: How matter becomes imagination. New York: Basic Books; 1995.
[14] Perlovsky LI, McManus MM. Maximum likelihood neural networks for sensor fusion and adaptive classification. Neural Networks 1991;4(1):89–102.
[15] Perlovsky LI. Mathematical concepts of intellect. In: Proc world congress on neural networks. San Diego, CA: Lawrence Erlbaum Associates; 1996. p. 1013–6.
[16] Aristotle IV BCE. In: Barnes J, editor. Complete works of Aristotle. Princeton, NJ: Princeton University Press; 1995.

[17] Aristotle IV BCE. Rhetoric for Alexander. In: Barnes J, editor. Complete works of Aristotle. Princeton, NJ: Princeton University Press; 1995.
[18] Aristotle IV BCE. Organon. In: Barnes J, editor. Complete works of Aristotle. Princeton, NJ: Princeton University Press; 1995. p. 18a28–19b4, 1011b24–1012a28.
[19] Hilbert D. The foundations of mathematics. In: van Heijenoort J, editor. From Frege to Gödel. Cambridge, MA: Harvard University Press; 1928/1967. p. 475.
[20] Gödel K. In: Feferman S, et al., editors. Kurt Gödel collected works, I. Oxford: Oxford University Press; 1986.
[21] Perlovsky LI. Conundrum of combinatorial complexity. IEEE Trans PAMI 1998;20(6):666–70.
[22] Bellman RE. Adaptive control processes. Princeton, NJ: Princeton University Press; 1961.
[23] Minsky ML. A framework for representing knowledge. In: Whinston PH, editor. The psychology of computer vision. New York: McGraw-Hill Book; 1975.
[24] Winston PH. Artificial intelligence. 2nd ed. Reading, MA: Addison-Wesley; 1984.
[25] A simple example of an adaptive model is linear regression: the knowledge is encapsulated in using linear model, and in the choice of variables, the uncertainty and adaptivity is in the unknown parameters, fitted to the data. Whereas linear regression uses one model, model-based systems used a large number of models. For example, a scene is described using geometric models of many objects. Parameters may include: size, orientation angles, color, illumination conditions, etc. A simple, still nontrivial problem, causing problems in applications till today, is tracking multiple objects in presence of clutter [see Ref. [26,27]].
[26] Singer RA, Sea RG, Housewright RB. Derivation and evaluation of improved tracking filters for use in dense multitarget environments. IEEE Trans Information Theory 1974;IT-20:423–32.
[27] Perlovsky LI, Webb VH, Bradley SR, Hansen CA. Improved ROTHR detection and tracking using MLANS. AGU Radio Sci 1998;33(4):1034–44.
[28] Perlovsky LI. Gödel theorem and semiotics. In: Proceedings of the conference on intelligent systems and semiotics ’96. Gaithersburg, MD, vol. 2; 1996. p. 14–8.
[29] Kecman V. Learning and soft computing: Support vector machines, neural networks, and fuzzy logic models (complex adaptive systems). Cambridge, MA: MIT Press; 2001.
[30] Marchal B. Theoretical computer science & the natural sciences. Phys Life Rev 2005;2(3):1–38.
[31] Penrose R. Shadows of the mind. Oxford: Oxford University Press; 1994.
[32] Catholic Encyclopedia. http://www.newadvent.org/cathen/08050b.htm.
[33] Wikipedia. http://en.wikipedia.org/wiki/Instinct.
[34] Jackendoff R. Foundations of language: Brain, meaning, grammar, evolution. New York: Oxford University Press; 2002.
[35] Searle J. Minds, brains, and programs. Cambridge: Cambridge University Press; 1980.
[36] Searle J. Intentionality: An essay in the philosophy of mind. Cambridge: Cambridge University Press; 1983.
[37] Searle’s views of intentionality are seen by many researchers as too narrow. For example, see [Workshop on Dynamics of Intentional Systems, 2005 International Joint Conference on Neural Networks, Montreal, Canada];
It is not possible to give in this review a complete treatment of such issues as intentionality and purposiveness, which differ in cognitive science, artificial intelligence, classical philosophy, and theory of action:
Davidson D. Essays on actions and events (philosophical essays of Donald Davidson). Oxford: Oxford University Press; 2002;
Bratman M. Faces of intention: Selected essays on intention and agency. Cambridge: Cambridge University Press; 1999;
Nevertheless, it is impossible to avoid these issues, because intentionality and purposiveness, as discussed later, are fundamental to living beings and to higher brain functions. Therefore, my approach in this review here and below is to use commonsense understanding of terms whenever possible, while noticing discrepancies among various understandings and to give corresponding references for further reading. I would like to emphasize again that the choice among various understandings of many philosophical, psychological, and cognitive terms used in this review is driven by four principles specified in the first section and is consistent with the mathematical theory presented in Section 5. Never a single narrow technical definition was selected to fit the mathematical structure. On the opposite, the mathematical structure turned out to be compatible with general understanding of these terms gradually developed since time of Socrates. The mutual compatibility of knowledge among Socrates, Plato, Aristotle, Kant, Jung, and contemporary research is emphasized, discrepancies are noticed, whenever relevant.
[38] Levine DS, Leven SJ, editors. Motivation, emotion, and goal direction in neural networks. Hillsdale, NJ: Erlbaum; 1992.
[39] Descartes, Rene; 1646. Passions of the Soul. Descartes deeply penetrated into the interaction between emotions and consciousness: “those who are the most excited by their passions are not those who knows them best and their passions are ... confused and obscure”. He did not differentiate thoughts and emotions consistently: “of all the various kinds of thoughts ... these passions”, “the action and the passion are thus always one and the same thing”. Descartes showed a bias toward unconscious and fused perception of emotions charateristical of a thinking psychological type.
[40] Kant I. Critique of judgment. London: Macmillan; 1790. Translated by J.H. Bernard in 1914.
[41] Sartre JP. Emotions, 1948;
See Sartre JP. Existentialism and human emotions, Reissue ed. Citadel Press; 1984.
[42] Jung CG. Psychological types. In: The collected works, vol. 6. Bollingen series, vol. XX. Princeton, NJ: Princeton University Press; 1921.
[43] Ortony A, Clore GL, Collins A. The cognitive structure of emotions. Cambridge: Cambridge University Press; 1990.
[44] LeDoux J. The emotional brain: The mysterious underpinnings of emotional life. New York: Simon & Schuster; 1998.
[45] Griffiths PE. What emotions really are: The problem of psychological categories. Chicago, IL: University of Chicago Press; 1998.
[46] Damasio AR. Descartes’ error: Emotion, reason, and the human brain. New York: Grosset/Putnam; 1994.
[47] Perlovsky LI. Cyberaesthetics: Aesthetics, learning, and control. In: STIS’98, Gaithersberg, MD; 1998;
Perlovsky LI. Emotions, learning, and control. In: Proc international symp intelligent control, intelligent systems & semiotics, Cambridge, MA; 1999. p. 131–7.

[48] Harlow HF, Mears C. The human model: Primate perspectives. Washington, DC: V.H. Winston and Sons; 1979.
[49] Berlyne DE. Conflict, arousal, and curiosity. New York: McGraw-Hill; 1960;
Berlyne DE. Pleasure, reward, preference: Their nature, determinants, and role in behavior. New York: Academic Press; 1973.
[50] Festinger L. A theory of cognitive dissonance. Stanford, CA: Stanford University Press; 1957.
[51] Clark A. The kludge in the machine. Mind and Language 1987;2.
[52] Minsky M. Smart machines. In: Brockman J, editor. The third culture. New York: Simon & Shuster; 1995. p. 152–66.
[53] Pinker S. Smart machines. In: Brockman J, editor. The third culture. New York: Simon & Shuster; 1995. p. 223–38.
[54] Chomsky N. New horizons in the study of language and mind. Cambridge, UK: Cambridge University Press; 2000.
[55] Perlovsky LI. Statistical limitations on molecular evolution. J Biomolecular Structure & Dynamics 2002;19(6):1031–43.
[56] Charness G, Levin D. When optimal choices feel wrong: A laboratory study of Bayesian updating, complexity, and affect. Department of Economics, UCSB. Departmental Working Papers. Paper 9-03. http://repositories.cdlib.org/ucsbecon/dwp/9-03.
[57] Carpenter GA, Grossberg S. A massively parallel architecture for a self-organizing neural pattern recognition machine. Computer Vision, Graphics and Image Processing 1987;37:54–115.
[58] Mathematically, the condition that the object h is present with 100% certainty, is expressed by normalization condition: $\int l(\mathbf{X}|h)\,\mathrm{d}\mathbf{X} = 1$. We should also mention another normalization condition: $\int l(\mathbf{X}(n))\,\mathrm{d}\mathbf{X}(n) = 1$, which expresses the fact that each received signal n is present with 100% certainty. This statement requires a qualification: it does not mean that the system is equally certain about all signals it receives, as discussed in Section 5.4, a lower level in the hierarchy is responsible for activating-sensing input signals for a higher level; the lower level "keeps track" of activation levels (certainties) of signals, which it sends up to the higher level.
[59] This construction is not unique. Expression (2) is inspired by the notion of likelihood between models and signals. Another general type of similarity measure suitable for the knowledge instinct, inspired the notion of mutual information in the models about the signals, is discussed in [8]. Here we would like to mention a modification of (2) for a specific case. Sometimes a set of observations, N, is more convenient to describe mathematically as a continuous flow of signals, for example, a flow of visual stimuli in time and space; then, it is convenient instead of Eq. (1) to consider its continuous version, $L = \exp \int_N \ln(\sum_{h \in H} r(h) l(\mathbf{X}(n)|h))$, where N is a continuum of bottom–up signals, such as in time-space.
[60] Perlovsky LI. Integrating language and cognition. IEEE Connections 2004;2(2):8–12.
[61] Perlovsky LI. Symbols: Integrated cognition and language. In: Loula A, Gudwin R, editors. Computational semiotics; 2006. In press.
[62] Perlovsky LI. Mathematical concepts of intellect. In: Proc. World Congress on Neural Networks. San Diego, CA: Lawrence Erlbaum Associates; 1996. p. 1013–6;
Perlovsky LI. Physical Concepts of intellect. Proc Russian Acad Sci 1997;354(3):320–3.
[63] A desirable property of convergence of model parameters to their correct values for objects they represent, or to the global maximum of the likelihood is not necessary in practical computer applications or in workings of the mind all the time. Perceptions are not always correct at a first sight, sometimes, we have to move, to change a point of view, before we unambiguously recognize an object. Same is true in computer applications. The practical aspects of efficient convergence are considered in [8,27,103].
[64] Cramer H. Mathematical methods of statistics. Princeton, NJ: Princeton University Press; 1946.
[65] Linnehan R, Mutz, Perlovsky LI, C, Weijers B, Schindler J, Brockett R. Detection of patterns below clutter in images. In: Int conf on integration of knowledge intensive multi-agent systems. Cambridge, MA; October 1–3, 2003.
[66] Zeki S. A vision of the brain. Oxford: Blackwell; 1993.
[67] See Ref. [40]. Some of my colleagues think that these are all Kantian ideas, other think that all is my invention and Kant never said anything that specific. I agree with both. I learned these ideas from Kant, but it is also true that other Kant readers did not understood Kant the same way. Every year there are dozens of papers published interpreting Kantian views, so clearly, the matter cannot be definitely settled within this review. I would just repeat that my inspiration for understanding Kant, Aristotle, along with many other philosophers and scientists is driven by a desire for a unified view of the mechanisms of the mind. I would emphasize, that I am skeptical about value of opposite approaches to understanding old texts: when I read that Kant or Aristotle did not understand this or that, usually I feel that the authors of these statements do not understand Kant or Aristotle. Similarly, I wonder, what Freud or Jung would think about their contemporary followers? In one movie comedy, Kurt Vonnegut wrote an essay for a college course about his own writing; this essay received C-. This joke reminds us to be modest about how well we can understand other people. I maintain that the only proof of correct understanding of any idea comes when it is positively integrated within the development of science.
[68] Deacon TW. The symbolic species: The co-evolution of language and the brain. New York: Norton; 1998.
[69] Peirce CS. Logic as semiotic: The theory of signs. In: Buchler K, editor. The Philosophical writing of C.S. Peirce (1893–1897). New York: Dover; 1955;
Peirce CS, et al. The essential Peirce: selected philosophical writings (1867–1893). In: Kloesel C, editor. Peirce Edition Project. Indiana University Press; 1992;
Hartshorne, Weiss, editors. Collected Papers vol. II, Elements of Logic. Cambridge, MA: Belknap; 1903.
[70] Russell B. Introduction to mathematical philosophy. London: George Allen and Unwin; 1919. p. 175 (note, a symbol for Russell is merely a sign).
[71] Carnap R. The logical syntax of language. Open Court Publishing Company; 1928/2003. Translated by A. Smeaton.
[72] Saussure FD. Course in general linguistics. New York: McGraw-Hill; 1916/1965. p. 98.
[73] Fiske J. Film, TV and the popular. In: Bell P, Hanet K, editors. Continuum. Australian J Media & Culture 1987;12:4–5.
[74] Derrida J. Writing and difference. Chicago: Chicago University Press; 1978.
[75] Chomsky N. Syntactic structures. Haag: Mouton; 1957.
[76] Chomsky N. Principles and parameters in syntactic theory. In: Hornstein N, Lightfoot D, editors. Explanation in linguistics. The logical problem of language acquisition. London: Longman; 1981.

[77] Chomsky N. The minimalist program. Cambridge, MA: MIT Press; 1995.
[78] Pinker S. The language instinct: How the mind creates language. Harper Perennial; 2000.
[79] Jackendoff RS. Semantics and cognition. Cambridge, MA: MIT Press; 1983.
[80] Lakoff G. Women, fire, and dangerous things. What categories reveal about the mind. Chicago: University of Chicago Press; 1987.
[81] Lakoff G, Johnson M. Metaphors we live by. Chicago: University of Chicago Press; 1980.
[82] Langacker RW. Foundations of cognitive grammar, vol. 1: Theoretical prerequisites. Stanford: Stanford University Press; 1987.
[83] Talmy L. The relation of grammar to cognition. In: Rudzka-Ostyn, editor. Topics in cognitive linguistics. Amsterdam: Benjamins; 1988. p. 165–205.
[84] R. Jackendoff works within both paradigms, nativist and cognitive, toward unifying both methods.
[85] Elman JL. Learning and development in neural networks: The importance of starting small. Cognition 1993;48:71–99.
[86] Elman JL. Finding structure in time. Cognitive Sci 1990;14:179–211.
[87] Perlovsky LI. Modeling field theory of higher cognitive functions. In: Loula A, Gudwin R, editors. Artificial cognition systems; 2006. Submitted for publication.
[88] Elman JL. Connectionist models of cognitive development: Where next? Trends Cognitive Sci. 2005;9(3):111–7.
[89] Olguin R, Tomasello M. Twenty-five-month-old children do not have a grammatical category of verb. Cognitive Development 1993;8:245–72.
[90] Tomasello M, Olguin R. Twenty-three-month-old children have a grammatical category of noun. Cognitive Development 1993;8:451–64.
[91] Nolfi S, Elman JL, Parisi D. Learning and evolution in neural networks. Adaptive Behavior 1994;3(1):5–28.
[92] Tomasello M. The cultural origins of human cognition. Cambridge, MA: Harvard University Press; 2001.
[93] Tomasello M. Constructing a language, a usage-based theory of language acquisition. Cambridge, MA: Harvard University Press; 2003.
[94] Rieger B, 2002. Private communication.
[95] Perlovsky LI. Computational linguistics, semiotics and the mind. In: Computational linguistics colloquium. University of Trier, 27 June 2002, Trier, Germany; 2002.
[96] Christiansen MH, Kirby S. Language evolution. New York: Oxford University Press; 2003.
[97] Chomsky N. The minimalist program. Cambridge, MA: MIT Press; 1995.
[98] Jackendoff R. Foundations of language: Brain, meaning, grammar, evolution. New York: Oxford University Press; 2002.
[99] Dawkins R. The selfish gene. Oxford: Oxford University Press; 1976.
[100] Nowak MA, Komarova NL, Niyogi P. Computational and evolutionary aspects of language. Nature 2002;417(6):611–7.
[101] Nowak MA, Krakauer DC. The evolution of language. Proc Natl Acad Sci USA 1999;96:8028–33. Evolution.
[102] Brighton, et al. Language as an evolutionary system. Phys Life Rev 2005;2(3).
[103] Fontanari JF, Perlovsky LI. Evolution of communication in a community of simple-minded agents. In: IEEE int conf on integration of knowledge intensive multi-agent sys. Waltham, MA; 2005.
[104] Fontanari JF, Perlovsky LI. Meaning creation and modeling field theory. In: IEEE int conf on integration of knowledge intensive multi-agent sys. Waltham, MA; 2005.
[105] Perlovsky LI, Fontanari JF. How language can guide intelligence. In: EVOLANG'06. Rome, Italy; 2006.
[106] Kant I. Critique of pure reason. New York: Willey Book; 1781. Translated by Meiklejohn JMD in 1943.
[107] Kant I. Critique of practical reason. Hafner; 1788. Translated by Bernard JH in 1986.
[108] Aristotle IV BCE. Metaphysics. In: Barnes J, editor. Complete works of Aristotle. Princeton, NJ: Princeton University Press; 1995. Translated by Ross WD.
[109] Plutarch II AD. Lives. Modern Library; 2001.
[110] Jaynes J. The origin of consciousness in the breakdown of the bicameral mind. 2nd ed. Boston, MA: Houghton Mifflin; 2000.
[111] Searle J. The rediscovery of the mind. Cambridge, MA: MIT Press; 1992.
[112] Searle JR. The mystery of consciousness. New York: New York Review of Books; 1997.
[113] Reyna VF, Brainerd CJ. Fuzzy-trace theory: An interim synthesis. Learning and Individual Differences 1995;7(1):1–75.
[114] Let me repeat, it is not possible to review all relevant points of view here, see [37].
[115] I'll repeat that Aristotle called intentionality the 'end causes of Forms' and he called the a priori contents of the mind the 'Formal causes' that is the mechanisms of Forms [Metaphysics].
[116] James W. The principles of psychology. New York: Dover Books; 1890, 1950.
[117] Consciousness of emotions and mechanisms of this type of consciousness is a separate, important, and interesting topic considered elsewhere [see [126]], it is beyond the scope of this review.
[118] Chalmers DJ. The conscious mind: In search of a fundamental theory. Oxford: Oxford University Press; 1997.
[119] Nagel T. What is it like to be a bat? Philos Rev 1974;11:207–12.
[120] Jung CG. Archetypes of the collective unconscious. In: The collected works, vol. 9, II. Bollingen series, vol. XX. Princeton, NJ: Princeton University Press; 1969.
[121] Taylor JG. Mind and consciousness: Towards a final answer? Phys Life Rev 2005;2(1):57.
[122] Perlovsky, L.I. Mathematics of intelligence and emotions. Invited plenary talk, 7th JCIS, Cary, NC, USA, 2003.
[123] Kant I. Anthropology from a pragmatic point of view. Boston, MA: Kluwer Academic; 1798. Translated by Gregor MJ in 1974.
[124] http://www.quotationspage.com/quote/26209.html.
[125] Perlovsky L. Aesthetics and mathematical theories of intellect. Iskusstvoznanie 2002;2(02):558–94 [in Russian].
[126] Perlovsky L. The knowledge instinct. New York: Basic Books; 2006.
[127] Peirce CS. Collected papers of Charles Sanders Peirce. Cambridge, MA: Harvard University Press; 1935–1966.
[128] Saussure FD. Course in general linguistics. New York: McGraw-Hill; 1916/1965. p. 98.

[129] Pribram K. Languages of the brain. Englewood Cliffs, NJ: Prentice-Hall; 1971.
[130] Fontanari JF, Perlovsky LI. Evolving compositionality in evolutionary language games. In: Evolution of languages, 6th int conf, Rome, Italy; 12–15 April, 2006.
[131] Feynman RP, Hibbs AR. Quantum mechanics and path integrals. New York: McGraw-Hill; 1965.
[132] Grossberg S. Linking mind to brain: The mathematics of biological intelligence. Notices Amer Math Soc 2000;47:1361–72.
[133] Westfall RS. Never at rest: A biography of Isaac Newton. Cambridge: Cambridge University Press; 1983.
[134] Koch C, Segev I, editors. Methods in neuronal modeling: From ions to networks. Cambridge, MA: MIT Press; 1998;
Hebb D. Organization of behavior. New York: J. Wiley & Sons; 1949.
[135] Piaget J. In: The Psychology of the Child. Basic Books; 2000. Translated by Weaver H.
[136] Grossberg S. Neural networks and natural intelligence. Cambridge, MA: MIT Press; 1988;
Zeki S. A vision of the brain. Oxford: Blackwell; 1993;
Freeman WJ. Mass action in the nervous system. New York: Academic Press; 1975.
[137] Koch C. The quest for consciousness: A neurobiological approach. Roberts & Company Publishers; 2004.
[138] Fontanari JF, Perlovsky LI. Evolution of communication in a community of simple-minded agents. In: IEEE int conf on integration of knowledge intensive multi-agent sys. Waltham, MA; 2005;
Fontanari JF, Perlovsky LI. Meaning creation and modeling field theory. In: IEEE int conf on integration of knowledge intensive multi-agent sys. Waltham, MA; 2005;
Fontanari JF, Perlovsky LI. Solvable null model for the distribution of word frequencies. Phys Rev E 2004;70:042901.
[139] Perlovsky LI. Neurodynamics and evolution of intentional communicating agents. In: Workshop on neurodynamics and intentional dynamical systems. International IEEE and INNS joint conference on neural networks (IJCNN'05). Montreal, Quebec, Canada; 2005;
Levine DS. Modeling the evolution of decision rules in the human brain. In: Workshop on neurodynamics and intentional dynamical systems. International IEEE and INNS joint conference on neural networks (IJCNN'05). Montreal, Quebec, Canada; 2005.