2019

Quantum metalanguage and the new cognitive synthesis

Alexey V. Melkikh
Ural Federal University

Andrei Khrennikov
Saint Petersburg National Research University of Information Technologies, Mechanics and Optics University ITMO

Roman V. Yampolskiy
University of Louisville, roman.yampolskiy@louisville.edu

Quantum Metalanguage and The New Cognitive Synthesis

Alexey V Melkikh[1]*, Andrei Khrennikov[2] and Roman V Yampolskiy[3]

ABSTRACT

Problems with mechanisms of thinking and cognition in many ways remain unresolved. Why are a priori inferences possible? Why can a human understand but a computer cannot? It has been shown that when creating new concepts, generalization is contradictory in the sense that to be created concepts must exist a priori, and therefore, they are not new. The process of knowledge acquisition is also contradictory, as it inevitably involves recognition, which can be realized only when there is an a priori standard. Known approaches of the framework of artificial intelligence (in particular, Bayesian) do not determine the origins of knowledge, as these approaches are effective only when "good" hypotheses are made. The formation of "good" hypotheses must occur a priori. To address these issues and paradoxes, a fundamentally new approach to problems of cognition that is based on completely innate behavioral programs is proposed. The process of cognition within the framework of the concept of a quantum metalanguage involves the selection of adequate a priori existing (innate) programs (logical variables and rules for working with them) that are most adequate to a given situation. The quantum properties of this metalanguage are necessary to implement such programs.

Key Words: Knowledge acquisition, Chinese Room, Understanding, Metalanguage, Quantum decision making, Generalization

DOI Number: 10.14704/nq.2019.17.01.1904 **NeuroQuantology 2019; 17(01):72-96**

Introduction

Problems of mechanisms of thinking in many ways remain unresolved. An answer to Kant's question "How can synthetic judgments be a priori?" posed 240 years ago has not yet been found. Indeed, despite significant progress made in the creation of artificial intelligence systems, these systems still cannot understand, acquire knowledge, create new concepts, etc. These properties are most fundamental to intellect and represent the greatest limitation of artificial systems. A number of authors believe that this challenge is not surmountable in principle and that a strong AI (completely coinciding with human intellect) will never be created.

Consideration of artificial intelligence in this regard is fundamentally important, as in this case we are dealing with systems, the structures of which we know. In regard to human thinking, the situation is different - we do not know exactly which structures store information and how such information is processed. This has given rise to the most important problem of what is innate and what is not.

Problems of cognition and thinking are at the intersection of different sciences (see for example, Miller, 2003) (Fig. 1).

However, the synthesis of these sciences in relation to problems of cognition has not yet taken place. Scientists of various cognitive sciences such as linguistics, epistemology, the philosophy of science, artificial intelligence, etc. do not often refer to one another's work. In particular, there is no single

Corresponding author: Alexey V Melkikh*

Address: [1]Ural Federal University, Yekaterinburg, 620002, Mira str. 19, Russia; 2International Center for Mathematical Modelling in Physics and Cognitive Sciences, Linnaeus University, Växjö, S-35195, Sweden, [2]National Research University of Information Technologies, Mechanics and Optics (ITMO), St. Petersburg 197101, Russia; [3]Computer Engineering and Computer Science Speed School of Engineering, University of Louisville, USA

e-mail ✉ melkikh2008@rambler.ru

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 October 2018**; Accepted:** 03 January 2019

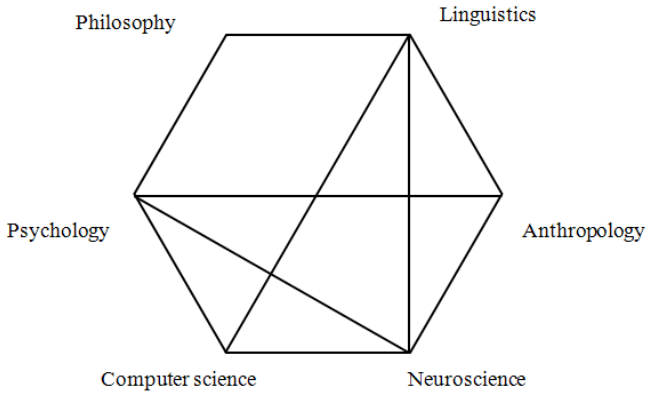

Figure 1. Cognition at the intersection of different sciences

approach to a major cognitive problem - the problem of knowledge acquisition. Thus, a new cognitive synthesis is needed.

Several approaches to the problem of cognition have been developed. A number of these approaches are based on the idea of the brain as a special structure that cannot be realized within the framework of any computing system. One objection to strong AI is Gödel's incompleteness theorem. A number of articles (see, for example, Penrose, 1994) state that a human understands, which is not provable, and that the difference between a human's thinking and a computer system is based on Gödel's incompleteness theorem, which concludes that thinking uses non-computable means. Another objection to strong AI is Searle's Chinese Room (1980), which is also related to the term understanding, with which, in the author's opinion, no formal artificial system is endowed. At the same time, the term understanding is poorly defined by the authors. Thus, a certain indeterminate property (or a set of properties such as the ability to acquire knowledge, to make generalizations, etc.) is attributed to a human, but proving anything under these conditions is difficult.

A potentially radical solution to the problem of acquiring knowledge is rooted in an assumption of innateness (partial or complete). Within the framework of the philosophy of science, this line of reasoning is connected to Fodor (1983). Chomsky, based on other considerations, came to a similar conclusion on innate language (Chomsky, 1957; 2014). However, this approach requires justification with the use of models and generalizations on thinking in the broadest sense of the word.

Some authors (see, for example, Aerts *et al*., 2011; Aerts *et al*., 2013; Pothos and Busemeyer, 2013; Khrennikov, 1999; Khrennikov, 2010a; Basieva *et al*., 2017) believe that a number of problems related to thinking can be solved within the framework of quantum mechanics. Korf, (2014) considered potential quantum effects on brain activity in a neurobiological context. The author considered three types of theories according to which the working of the brain is related to quantum mechanics. The first states that quantum mechanics are involved in the behaviors of molecules but are not related to higher brain functions or consciousness. The second states that the brain is viewed as part of a world based on quantum mechanics (Hameroff and Penrose, 2014, Smith, 2006; 2009). The third uses quantum mechanics to model mental processes and can be classified as quantum-like, as the authors view quantum mechanics as a convenient formalism for solving a number of decision-making problems.

One of the motivations of applying quantum mechanics to processes of thinking was experiments on decision making. Such experiments examine, for example, the Elsberg paradox, the breaking of the sure-thing principle, and other phenomena. In (de Barros, 2012; Khrennikov, 2011; Aerts *et al*., 2011; Basti *et al*., 2017) it was shown that the formalism of quantum mechanics can help explain these paradoxes. However, a question remains as to how this quantum(-like) behavior is realized. Which structures of the brain are responsible for it?

Thus, quantum mechanics has proven useful for modeling the workings of cognitive functions. The purpose of this article is to go further and show that quantum mechanics can be used to solve the most fundamental problems of thinking, such as those of knowledge acquisition, understanding, new concept generation, etc. The very phenomenon of thinking is systemic in nature, and thus a systems approach must be applied to problems of thinking in general and to problems of acquiring knowledge in particular.

The problem of human consciousness is not considered in this work. This, while related to the problem of knowledge acquisition, is another matter to which the great literature is devoted (Chalmers, 1996; Dennett, 2005).

Searle's Chinese room, objections to strong AI and the problem of knowledge acquisition

Central to the problem of cognition is understanding.

There are two different aspects of understanding. The first is related to adequate actions that show that a human (the intellectual

system) really understands what is in front of him or her. The second is related to consciousness and to feelings about what one sees. Feelings (qualia) are thus attributed only to humans and not to an arbitrary intellectual system. In this article we only consider the first aspect of understanding. It is this aspect that has a direct bearing on the problem of knowledge acquisition, which is the main focus of this article. For the second aspect, research on it has a long history and several articles have been devoted to it (see for example, Yampolskiy, 2017). Studying and modeling this second aspect is difficult because it is difficult to measure in an objective and human-independent manner.

By understanding we of course refer to an understanding of something new. Otherwise we are considering an already known object, in relation to which we can act adequately, i.e., its understanding is a priori present.

One of the key thought experiments conducted on this topic was Searle's Chinese Room experiment.

In 1980, Searle (1980), with the help of a thought experiment entitled "The Chinese Room", put forward a serious objection to the existence of strong AI. In fact, this objection is critical to understanding thinking as such. In the mental experiment, a human in a room could communicate with the outside world only with instructions written in Chinese, which he did not know. Comments made on the instructions were written in English. However, the instructions were designed in such a way that they could be executed with the help of comments but without understanding what exactly is being done.

As a result, Searle concludes that while such a system can pass a Turing test, no understanding of the language of the system is given, meaning that the Turing test is not an adequate test of cognitive abilities. Searle's arguments are aimed at criticizing the view of "strong" artificial intelligence, according to which computers equipped with a corresponding program can understand natural language and possess other mental abilities peculiar to people. The hypothesis of weak artificial intelligence only suggests that computers are capable of imitating a human's mental abilities, which corresponds to the mental experiments conducted by Searle.

As Searle noted, syntax is not sufficient for semantics.

The problem here is that within the framework of this thought experiment the term understanding is poorly defined. The author believes that this is something that is taken for granted. However, the absence of a definition given explicitly does not allow us to conclude whether an artificial intelligence system is able to understand or whether it merely imitates this process.

A large number of articles have been published on the problem of the Chinese Room (see for example Chalmers, 1996; Barsalou, 1999; Bishop *et al*., 2013). Quantum versions of the Chinese Room have also been developed. For example, Maruyama (2016) considered another argument of Searle based upon the observer-relativity of computation. According to the author, calculations are dependent on the observer but human intellect is not. Therefore, intellect cannot be simulated through calculations. Calculations must instead be interpreted by someone (that is, understood). The author considered the quantum mechanics of categories, from which quantum linguistics arose.

A definition of understanding can be given based on experiments that determine whether a system (human) behaves appropriately after receiving information from outside or not. What does it mean, for example, to understand the term "electrical socket"? It means that a human can use this tool to perform certain actions (turn on a table lamp, etc.). If this term is written, for example, in Vietnamese, which the person does not know, then the actions of the human will not be adequate, i.e., actions are given no value as a result of receiving such a message.

Here, it is necessary to note the importance of the recognition procedure, which has typically received little attention in reference to the context of the problem under consideration. This procedure in its most general form involves the comparison of an image obtained by receptors to some standard. As a result, two options may be available: the image may coincide with the standard (correspond to it) or not. For the first case, it is known what the image is, as a priori programs work with this standard and generate adequate results, etc. Rather, the first case corresponds to the achievement of understanding. However, in this case we do not understand something new. In the second case, when an image is not recognized, no a priori programs are involved, adequate behavior is not realized, and no understanding is achieved. This is a general procedure (which underlies large subject - pattern recognition) that is not directly related to

image processing methods, information transferring, physical media, etc. That is, such a procedure should be used in both living and artificial systems. If the Chinese language is not determined based on standards within an artificial intellectual system, then it cannot act adequately in this regard, i.e., the system does not understand it.

When a human is able to understand (as manifested in adequate actions) what is written in Chinese (never having studied it before), then this is possible only in one version - when this language nevertheless is a priori embedded in our consciousness as a standard. The application of this possibility is discussed in Section 7.

Other objections to strong AI have been put forward by Penrose (1994). Penrose also believes that intellect is an understanding that involves consciousness. In the opinion of the author, a distinction must be made between true intellect and attempts to model it. Penrose argues that understanding is not computable. As proof, he uses Gödel's incompleteness theorem and concludes that conscious thinking has a non-computable component. The author, however, concludes that when the ability to understand is non-algorithmic, then it can be simply and easily explained through natural selection. As a means to solve this problem Penrose suggested that quantum events occurring in the microtubules of neurons are not computable.

Here, we would like to note that understanding does not have to involve consciousness. When an intelligent system of an arbitrary nature is capable of adequately responding to a certain set of symbols, we must conclude that it understands them. To ensure that it is conscious of them is quite difficult from an experimental point of view. The very concept of "consciousness" does not have a clear definition.

If understanding is considered in this context (as an adequate behavior), then different forms of understanding must occur at the molecular level (including neurons). Any cell, in addition to transforming and transporting substances, at the same time processes information. For example, in (Gorlich *et al*., 2011) the cell is considered a semantic system. The authors consider the codes that organize the semantic dimension of molecular information. Such operations implicitly involve recognition. If, for example, a membrane receptor recognizes an external signal (i.e., a reaction occurs between it and a certain molecule according to the key-lock principle), this results in the further processing of this signal, in its amplification, etc. We can state that the cell has "understood" the signal. This understanding is expressed as the cell's adequate response to a certain change in the environment. However, in this case, the signal cannot be new for the cell or for an arbitrary system.

Thus, on one hand, the argument of the Chinese Room is important in the sense that it is possible to show through this framework that understanding (and as will be shown below its consequences such as the formation of new concepts, abstractions, etc.) cannot be obtained based on any formal procedures or programs developed accordingly. On the other hand, the Chinese Room itself is nothing special relative to other situations, as a human perceives the world around him or her only through receptors (senses). In this sense, we are all partially in the Chinese Room. How then do we understand anything new? If a child has learned a language, how does he manage to understand it? The assertions of many authors that a child is taught by trial and error cannot be accepted as satisfactory, as a computer can also act in such a way. In this case, an equally important question arises as to how many trials and errors are required? This question is considered below.

Objections like the Chinese Room are closely related to the problem of knowledge acquisition. As it was noted above, when a human or an intellectual system recognizes an environment with the help of receptors, only two possibilities are realized. 1 - The image that appears in the environment is not recognized, and as a result no a priori programs for working with it are available. As a result, behaviors will not become more adequate after the registration of this image. 2 - The image is recognized, but this can occur only when the system is given a priori standard for its recognition. In this case, the system's behavior can become more adequate through the a priori application of behavior programs. However, knowledge acquisition in this case also does not occur, as programs of behavior existed in the system a priori. This problem is considered by Melkikh, (2014a). In particular, in (Melkikh, 2014a) two classes of behavior (Table 1) are distinguished for animals and arbitrary intellectual systems:

It is shown that the behavior of type B is contradictory and cannot be realized. According to (Melkikh, 2014a), such a system of recognition and decision-making can be described using the following model:

Table 1. Two classes of behavior are distinguished for animals and arbitrary intellectual systems (Melkikh 2014a).

Behavior based on innate programs (A)	**Behavior based on knowledge acquisition (B)**
Includes conditional and unconditional reflexes and the simplest forms of learning.	Includes learning experiences (interactions with other animals) and the logical formulation of new concepts.
An animal (intelligent system) meets a recognized object. This is accompanied by the running of innate programs. An unrecognized object can only trigger innate programs. An adequate response to the unrecognized object is impossible.	An animal (intelligent system) meets an unrecognized object and begins to respond to it adequately through the acquisition of knowledge.

Property 1: The system contains internal independent objects (words of the language) $Q_1 \dots Q_n$. Operations $P_1 \dots P_m$ on objects are possible. Operations of new object (operation) creation are not defined.

Property 2: The system includes a receptor that receives information on the state of the environment. The receptor is capable of detecting objects $q_1 \dots q_m$. In this case, operations $p_1 \dots p_k$ can be performed on objects. The receiving signal is compared to internal objects present in the system (pattern recognition). The system compares internal and external objects of the language. Mathematically, this comparison can be expressed in the form of equivalence (equality) of any internal object to some of the external objects.

Property 3: From recognition (matching), some internal object operations trigger effectors; the system acts on the external environment. Every action is evaluated using function $\Phi_i(Q_i, P_i)$ (gain). This value is greater when the system is more successful. Without a loss of generality, we can assume that the system is designed such that there is a maximum condition: $\Phi_i(Q_i, P_i) \to \max$. The term of "equivalence" is closely related to the gain and represents an adequate reflection of reality.

Property 4: The system has limited memory storage, which can store the results of previous measurements or of intermediate operations.

The pattern recognition and decision-making scheme can then be represented as the following form: (Fig.2)

Let image Ω appear to the organism (system) with properties 1-4, which can be described in terms of the external language of receptors (registered). However, no objects of the internal language are relevant (in terms of equivalence) to external language objects corresponding to this image. Then,

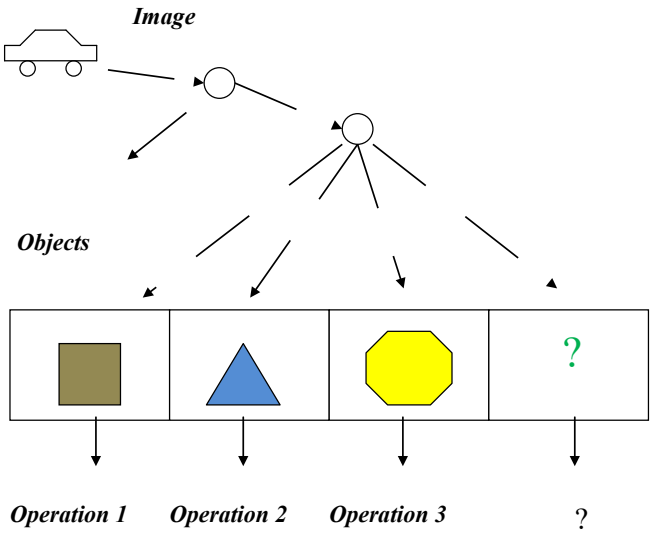

Figure 2. The pattern recognition and decision-making scheme

the effector cannot perform operations or activities that lead the system to obtain a certain gain (payoff).

Consider a universal translator that translates one language into another and that can work with an unrecognized image. In this case, the algorithm cannot be run because the conditions required to run it, cannot be formulated.

Suppose that the operation of supplementing a list of operations and objects from external alphabet phrases exists. However, this operation is controversial because it can be only performed if an external alphabet is part of the internal alphabet. When a phrase is part of an internal alphabet, the registered object is recognized, contradicting the initial claim. In this context, language problems require separate consideration and are discussed below in paragraph 5.

In regard to pattern recognition and decision-making, the next version of the statement is controversial: "to create a new standard requires a new standard." When an image does not match an internal standard, the system must create a new standard to make an adequate decision. However, its creation is self-contradictory because the algorithm that creates such standards must know them in advance.

The contradictory nature of the acquisition of knowledge based on learning automata is represented by the following diagram (Melkikh, 2014a) (Fig. 3):

Thus, it is not clear how an arbitrary system acquires knowledge, as new information is not valuable, and valuable one is not new. That is, when

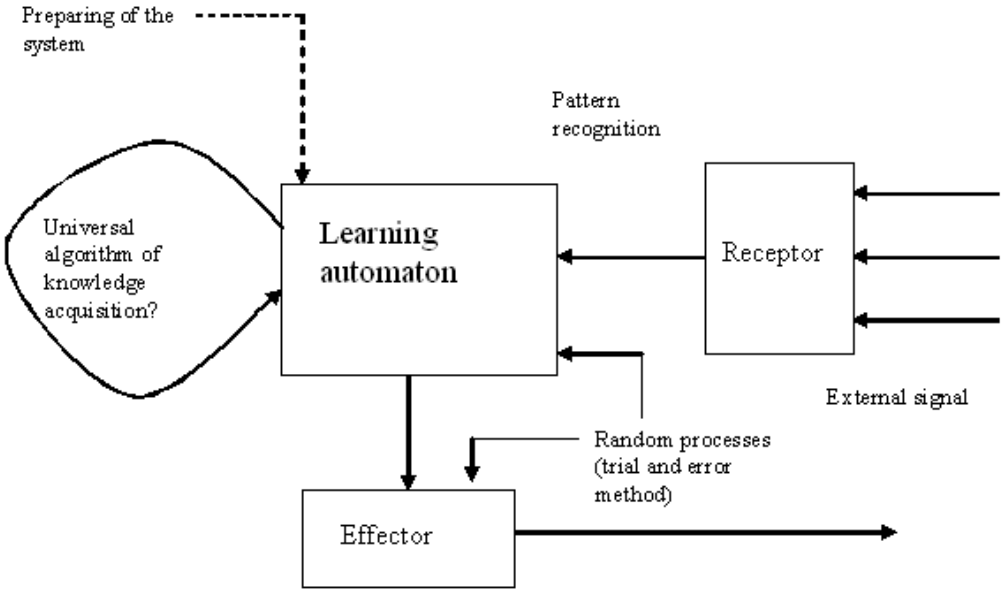

Figure 3. The contradictory nature of knowledge acquisition and learning automaton

a human is able to understand something that is not available to a computer, then he is capable of acquiring knowledge.

Often, the acquisition of knowledge is associated with the concept of "intelligence." There are many definitions of "intelligence" (see, for example, Legg and Hutter, 2007). For example, Gottfredson (1997) defines this concept as:

«A very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience. It is not merely book learning, a narrow academic skill, or test-taking smarts. Rather, it reflects a broader and deeper capability for comprehending our surroundings—"catching on," "making sense" of things, or "figuring out" what to do».

Thus, the problem of knowledge acquisition and the problem of understanding are closely related. To determine whether an arbitrary system is able to understand or acquire knowledge, it is necessary to explicitly take the recognition process into account.

Ascending and descending approaches to AI and the problem of knowledge acquisition

In reference to the problems discussed above it is important to consider artificial intelligence systems, which are distinguished (from the brain) by the fact that their structures are known.

In artificial intelligence, two basic approaches are used: ascending and descending approaches. An ascending approach involves the use of neural networks, including promising newly developed networks of deep learning (see for example, LeCun *et al*., 2015; Sze *et al*., 2017). To train a neural network on something, we must explore a large collection of preliminary examples. Note that in these examples, someone (a human) must explicitly indicate what kind of thing it is. This instruction occurs in a special mode - the training mode. The neural network itself does not need to understand anything - it is understood by the one who teaches it, i.e., the human. In the case of unsupervised learning, neural networks are only capable of clustering information. In this case, there is no recognition.

Recently, the AlphaGo Zero algorithm achieved superhuman performance in the game of Go, Chess and other games using deep convolutional neural networks, trained solely by reinforcement learning from games of self-play (see, for example, Silver *et al*., 2017). It was demonstrated that a general-purpose reinforcement learning algorithm can achieve, tabula rasa, (without any additional domain knowledge except the rules of the game) superhuman performance across many challenging domains. In this case, however, new images do not arise, and there is no understanding.

Note also that machine learning is applicable only to problems for which there are sufficient initial data (see for example Abu-Mostafa *et al*., 2012). Machine learning does not create information but rather uses it based on already available data. When there are not enough training data, machine learning

methods do not work. It is also necessary to define the term "data." Clearly, if these data do not correspond to the "meaning" a priori, it will not appear as a result of further training. The collation of meaning to a signal is a human's task.

The descending approach is related to the application of the Bayesian formula for conditional probabilities and generally theory of Bayesian networks. Hypotheses (generative models) are put forward and subsequently verified (see for example, Lake *et al*., 2015; Bonawitz *et al*., 2014). A Bayesian system can be trained on a small number of examples, but this requires the use of a set of hypotheses that the system itself cannot offer. Hypotheses are again formulated by those who understand (a human). This method is considered in more detail in the following section.

Several works have been devoted to children's learning processes. A number of authors believe (see for example, Bonawitz *et al*., 2014) that children learn in a Bayesian manner. However, this raises question of who in this case forms hypotheses? When a hypothesis is formed by an external system (another human), then how does the trainee understand what is being said when first presented with this situation (new objects or images)? In the second case it can be assumed that hypotheses are congenital (i.e., they exist prior to the learning process).

It is often believed that children learn from examples by default, but this is not the case. For example, in (Melkikh, 2014a) social learning was considered. Social learning is often considered separately owing to its complexity and importance to complex communities (transfer of experience from parents to descendants and so on). However, in terms of the problem addressed in this article, social learning is not fundamentally different from other forms of learning. We consider a case in which an organism receives new information from its parents or from other animals. Can the organism adapt to new conditions in this case? Given what has been said above, the answer is no.

Importantly we note that in this case, the organism also receives information via receptors. Therefore, any other organism (a member of the flock, the parent, etc.) is interpreted as part of the environment. Consequently, the aforementioned scheme of measurement - recognition – decision making holds. Regardless of actions the parent takes, they can be adequately interpreted only in one case: when a signal has been recognized (compared to a standard sample). In turn, this comparison is only possible when a standard sample has already been made available to the organism (as congenital).

Note also that "supervised learning" in animals and humans is fundamentally different from "supervised learning" in neural networks, in that a neural network "teacher" can prepare a system (i.e., to set connection weights between neurons form outside), while for animals this is impossible. All "teachers'" acts can be perceived only through pattern recognition.

Thus, a child can "learn" only what he is ready to absorb, for which he already has all necessary programs, and training is only available to launch such programs in a timely manner and to launch programs that are needed in the near future.

In (Melkikh, 2014a) basic methods of solving problems in AI are considered and it is shown that all of these methods are based on the use of a priori information. In addition to neural networks we list the following tools:

1. Heuristic methods of problem solving are based on aprioristic information of the object domain that comprises a problem. Heuristics may prove to be incapable of finding a solution altogether. This limitation cannot be removed even from the best heuristics (Garey and Johnson, 1978). A key question concerns how heuristics arise. The theory does not address questions of how we can develop a new heuristic.
2. The recursive search represents a natural means of realizing such strategies of artificial intelligence through a graph search. However, an exact objective should be set for this method to be used. When objects (or even one object) are not defined, a recursive procedure cannot be applied (Luger, 2003).
3. Production systems are used for conflict resolution. In this case, the system requires an a priori standard sample (pattern) that determines the possibility of using the rules of production systems (Luger, 1994).
4. Expert systems are based on the same rules (Luger 2003). The core of an expert system is a knowledge base that includes knowledge generated from a particular applied domain. Knowledge in an object domain determines and updates an expert

database (Minsky, 1987). Can we possibly imagine a system that acts as an expert for itself? This is clearly not possible because a code can be corrected only when correct knowledge is known.

5. Recently, hybrid connectionist-symbolic systems based on a synthesis of neural networks and based on symbolic representations of knowledge (see for example, Vilhelm *et al*., 2000; Kamsu-Foguem *et al*., 2012; d'Avila Garcez *et al*., 2009) have been proposed. These systems facilitate knowledge acquisition through a variety of applications. However, the term "knowledge acquisition" is defined for such a system in a different sense than it is in this article. In the context of this approach, "knowledge acquisition" denotes that with help of knowledge, some useful tasks can be completed (e.g., in medicine). The functioning of such systems a priori assumes that all characters (or simple signals) or words written in a specific language are recognized by a human or computer. In this sense, they are innate. In this case, problems related to determining whether knowledge is innate to humans are not considered.

We can also consider a hybrid variant: interactions between the brain and machine. Indeed, at the moment, control technologies that use brain signals to control a computer cursor or hand prosthesis are actively being developed (see for example, Mathot *et al*., 2016). However, these technologies do not relate to the term understanding.

Thus, an analysis of main approaches to the problem of knowledge acquisition in artificial intelligence systems leads to the conclusion that in either case all knowledge is a priori. In the first case, a teacher (who can manage the system in a special mode) has knowledge and in the second, the system itself has them. No new knowledge appears in this case.

In (Melkikh, 2014a) it is noted that a special regime of system preparation can exist. As changes within an organism and its interactions with its environment are crucial when considering knowledge acquisition and learning, these interactions should be defined more precisely. Regarding processes of knowledge acquisition and learning, interactions between an organism and its environment must be performed only through recognition. The organism obtains all information on its environment through its senses. Furthermore, this information is subject to recognition, after which the organism decides what to do in the current situation. However, in principle, an organism can interact with its environment in another way – through preparation. The structure of the system can be directly changed by acting on it from the outside (not through the senses). This method is widely used in engineering systems. For example, a computer user can replace any part of a system to create a different system. However, this method is not related to learning or knowledge acquisition (as in this case another system learns); in addition, the method is not used in living systems. Thus, the system's preparation is not considered further.

A possible answer to questions on the origins of knowledge in artificial intelligence systems concerns the existence of universal solvers. For example, an analysis of universal solvers (Hernandez-Orallo, 2016) shows that

1. Some authors argue that there is an effective search method, as the interval in which a number (ε-acceptable policy) is located is known. Additionally, various search variants are considered. In principle, however, these search methods implicitly assume that the object under study is recognized. When it is not recognized (it is new), the interval in which one wishes to search is not known and the search space does not offer any potential values.
2. The author (Hernandez-Orallo, 2016) appeals to evolution as a search method, as by default it is assumed that its opportunities to solve problems are unlimited. However, references to evolution are untenable, as we do not know how a search actually occurs. The assumption of the Darwinian search mechanism is presented with serious conceptual difficulties associated with a need to search through a large number of variants (see for example Melkikh, 2014c; Melkikh and Khrennikov, 2017). The fact that this or that adaptation occurs in life does not say anything about the fact that a universal solver is used. Rather, it indicates that a system has a priori information on the solution which should be found.
3. In this connection, a question arises as to whether there is at least one example of an artificial universal solver that would work in life (with real objects). The practice of solving technical problems via intelligent systems shows that there are no such solvers.

4. Recognition models appear to be fundamentally important. This process cannot be described through words alone and must be modeled.

Can an intellectual system self-improve while solving problems? If we could create a form of superintelligence that could improve itself, then this would be a breakthrough in the field of cognitive sciences. However, it is important to accurately define concepts. An improvement made in one area does not imply anything about improving work in another area. In (Yampolskiy, 2014; 2015), various forms of intellectual system self-improvement are considered:

- self-modification (1),

- self-improvement (self-adaptation) (2),

- recursive self-improvement (3).

The purpose of the first method is to protect the program itself and not to improve it.

The second concerns the goal of many software products to adjust the system to the user. For example, genetic algorithms optimize parameters for the analysis of a particular system through various learning algorithms such as meta-calculations.

The third allows one to replace an algorithm with a new algorithm. The author notes that this approach has not yet been employed. In this regard, it should be noted that this third method of self-improvement cannot be implemented in principle, as an algorithm must replace the algorithm that the new algorithm must contain a priori.

Bayes hypotheses and trial and error method

As noted above, the trial and error method is considered a common method of learning used by humans and animals. In some cases, a human needs to make only a few attempts to learn something. However, why does such a method work? Let us show that the working capacity of the trial and error method is closely related to Bayesian hypotheses that underlie it. We write the Bayesian formula as follows:

$$P(A/B)=\frac{P(A)P(B/A)}{P(B)}, \qquad (3)$$

Under A and B we have the following:

P (A) is the a priori probability of hypothesis A;

P (A / B) is the probability of hypothesis A occurring with the occurrence of event B (a posteriori probability);

P (B / A) is the probability of event B occurring with the truth of hypothesis A;

P (B) is the probability of event B occurring.

If the hypothesis put forward (explicitly or implicitly) is valid (good, corresponds to reality), then probability P (A / B) should be high. This means that learning will follow

$$n\sim\frac{1}{P(A/B)},$$

steps, which is a small number. However, questions then arise as to where this hypothesis comes from. Why is this hypothesis valid in any sense? As a natural and non-contradictory answer to this question, the hypotheses themselves are congenital or are derived from inborn knowledge based on rules of inference. Trials and errors in this case are only a consequence of a lack of such knowledge due to uncertainties related to the environment. Information obtained through experiments and the number of experiments the must be conducted to obtain such information are related (see appendix 1).

A priori information can be calculated from the Shannon formula based on a priori probability. Thus, a good hypothesis is a hypothesis that contains a considerable amount of a priori information on the learning object. This information cannot arise through the learning process but should be a priori, i.e., congenital. However, this means that knowledge is not acquired through such a process and only applies a priori knowledge most appropriate to the given situation. When a human (animal) immediately understands what is before him and which actions must be taken without trial and error (insight), then this is possible only when the human already knew this in advance (i.e., enough a priori information was available on the object). Whether a human perceives this process or not is another matter.

If, as a result of using the trial and error method, the results of these trials (experiments) accumulate, then repetition can be applied to reduce errors. In (Melkikh, 2014a) the model of errors at recognition is considered. According to (Melkikh, 2014a), errors occur each time a device measures its environment, potentially leading to erroneous recognition. Therefore, the larger the degree of measurement error, the smaller the fitness Φ value (fitness is at a maximum when errors are absent). We refer to the maximum fitness level as Φ_0. Then we can write the following (when errors are relatively small):

$$\Phi = \Phi_0 - \Delta,$$

where Δ is a positive value denoting a decrease in fitness resulting from a measurement error. From error theory we know that random error decreases with repeated measurements as the number of measurements increases (when systematic inaccuracy is present, it can be included in the maximum fitness value). However, each measurement requires energy to complete and as a result fitness values will decrease. We denote a decrease in fitness resulting from energy expended for a single measurement as ε. Then we have the following formula for the fitness value:

$$\Phi = \Phi_0 - \varepsilon n - \frac{\Delta}{\sqrt{n}}, \quad (4)$$

where n is the number of measurements.

Factor $1/\sqrt{n}$ appears with a decrease in the random error value occurring from repeated measurements. This formula can be used when the measurement time is short compared to the characteristic behavior time. Alternatively, when measurements are collected slowly, the fitness value will decrease because the organism will not have enough time to measure rapid changes in the environment.

The fitness value (4) may have an extremum with respect to n, and thus we have:

$$n = \left(\frac{\Delta}{2\varepsilon}\right)^{2/3},$$

$$\Phi(X,Y) = \Phi_0(X,Y) - \varepsilon^{1/3}\Delta^{1/3}\left(\frac{1}{2^{2/3}} + 2^{1/3}\right).$$

As the number of measurements is an integer, the extremum exists only at Δ/ε > 2. Therefore, an organism, which occupies a complex environment, may find it unfavorable to immediately change its behavior as its environment is altered and instead might prefer to perform a set of measurements and only then begin to act. In other words, to initiate available programs an organism should first recognize its environment and identify necessary programs.

This behavior explains the fact that living organisms form conditioned reflexes most often as a result of the recurrence of an external signal. Thus, repeated measurements of the state of the environment from an organism decrease errors during the operation of aprioristic programs.

What is innate?

To understand mechanisms of acquiring knowledge it is important to clarify the question of what can be attributed to the innate in living and artificial systems.

What is innate? The answer to this question is quite complex. Even within the framework of cognitive sciences, no unambiguous definition of innate is given.

By the term instinct, as applied to the behavior of animals and humans, we refer to complex stereotyped forms of activity that are inherent of all individuals of this species and that are inherited and do not require training. Instinct represents a form of species memory passed down from generation to generation by inheritance. The behaviors of animals and humans involve a complex interlacing of innate and acquired skills.

Ethologists define instincts as specialized morphological structures (a temporary organ of the animal) that naturally appear in a specific social situation (Lorenz, 1950). An instinctive reaction is realized automatically with any presentation of specific stimuli and is not corrected by circumstances of the given context or by the animal's past experiences. In this case, innate response circuits are triggered.

For example, Samuels (2004) considers several definitions of innateness. The author notes that the theory of innateness has been used to explain many psychic phenomena. However, the concept should be better defined. The author considers several definitions of innateness:

- Innateness as non-acquisition. Is universal grammar innate? However, non-acquisition is also poorly defined. Minimum definition is: a characteristic is acquired from an object (e.g., an organism) when and only when there is a period of time in which an object has the characteristic in question and a prior period in which it does not. However, this is not enough, as at development all cognitive abilities are absent.
- Innateness as a presence at birth. That is they are acquired in the minimal sense - during growth. However, it is possible to train prior to birth. Sexual signs are also absent (in part) at birth.
- Innateness as a consequence of internal causes.

— Innateness as genetic determinism. However, here it is impossible to trace a clear causal relationship.

— Innateness as an invariant of development.

— Innateness as something that has not been taught. We in turn require a clear definition of training.

There are two prior opposing theories that link behavior to genotypes. One theory conditionally referred to as "behaviorism" denies the importance of genes. However, another ("genetic determinism") states that genes completely determine behavior.

At present these extreme views have not been confirmed. Studies of the individual development of organisms show that in different environments the same genotype can be expressed differently, and the phenotypes and behaviors of animals and humans depend not only on genes but also on the environment (see for example Argawal, 2001; Whitman and Agrawal, 2009).

For example, the study of the IQs of monozygotic twins (see for example, Butcher *et al*., 2008) unambiguously shows that genes play an important role in intelligence. However, as shown by experiments, identifying genes responsible for specific aspects of behavior (e.g., learning) is difficult even for the simplest organisms. For example, Glanzman (2010) notes that the presence of molecular pathways is not sufficient to understand learning. Examples of Aplisia and Drosophila behavior show that molecules such as GABA and glutamate are associated with neuronal plasticity. However, by themselves molecules cannot explain behaviors without consideration of neural networks. «The take-home lesson from the two studies discussed here is that knowledge of the key molecular players does not provide a short cut to understanding memory and cognition» (Glanzman, 2010). This conclusion can be attributed to other organisms and behaviors.

According to experiments on human genome mapping, a very small number of genetic markers are related to intelligence test results (see, for example, Coleman *et al*., 2018; Hill *et al*., 2018).

Thus, an organism can change its phenotype in response to changes in its environment (phenotypic plasticity). If we consider behavior, the special case of phenotypic plasticity can be considered as a form of neuroplasticity, as rooted in changes in connections between neurons and in the strength of synaptic contacts (synaptic plasticity) influenced by experience. Many mechanisms of synaptic plasticity are common among vertebrates and invertebrates (see for example, Glanzman, 2010).

Noble (2006) examined the relationship between different levels of causality in cells. In his view, genes do not represent a program because many properties of an organism (cells) can be realized only at the higher systems level. However, mechanisms of this higher level remain unclear. Does this level involve the use of algorithms? If so, where do they come from, and if not, how does it work?

Various researchers (for example, Fodor, Pask, Osherson) have suggested that some concepts of human language are congenital. In this case, however, formation mechanisms of remaining concepts remain unclear.

For example, Fodor (1983) suggested that there is a special language of thinking referred to as mentaliz. According to Fodor, mentaliz is unique in that it cannot be learned from one's native language, as it is innate. This applies, in his opinion, to a number of other concepts (e.g., BACHELOR, EFFECT, ISLAND, TRAPEZOID, WEEK) that the author attributes to primitive and innate. It remains unclear, however, how other concepts arise.

Fodor and Pylyshyn (2016) explored an analogy between thinking and computers based on computation: "Cognitive processes are computations, which is to say that they apply to ...." This analogy also means that cognitive processes must be based on innate knowledge, as computers work this way.

Chomsky's theory (see, for example, Chomsky, 1957; 1965) posits that language involves the use of deep and surface structures. Surface structures 'face outward' as represented by spoken utterances while deep structures 'face inward' and express underlying relations between words and conceptual meanings. Transformational grammar is generative grammar that applies a limited series of rules expressed in mathematical notation that transforms deep structures into well-formed surface structures. Transformational grammar thus relates meaning and sound.

The Chomsky hierarchy is a hierarchy of classes of formal grammar. The hierarchy imposes a logical structure to different language classes and provides a basis for understanding the relationship between grammars (devices that enumerate valid sentences within languages). In order of increasing

expressive power, it considers regular, context-free, context-sensitive, and recursively enumerable grammars. Each class represents a strict subset of the class above it, i.e., each successive class can generate a broader set of formal languages.

Later versions of Chomsky's theory (the minimalist program, Chomsky, 2015) make strong statements on universal grammar. According to his view, grammatical principles underlying languages are innate and unchanged, and differences between languages of the world can be explained in terms of parametric brain settings that can be compared to switches. According to this point of view, a child must learn lexical units (i.e., words) and morphemes to learn a language and must determine necessary parameter values in reference to several key examples.

This approach, according to Chomsky, explains the amazing speed at which children can learn languages, consistent stages of child language learning regardless of languages learned, and characteristic errors that children make while learning their native language and seemingly logical errors that are not made. According to Chomsky, the absence or occurrence of such errors is related to the method used: generalized (innate) or dependent on a specific language.

Chomsky notes that the infinite number of proposals that a human can make serves as a strong reason to reject the behaviorist concept of language teaching by reinforcing (fixing) conditioned reflexes. Young children can make new proposals not backed by past behavioral experience. The understanding of language is conditioned less by past experiences than by the so-called Language Acquisition Device, the internal structure of the human psyche. The mechanism of language learning determines the volume of permissible grammatical constructions and helps a child master new grammatical constructions from the language he has heard.

According to the author, congenital grammars arose through random mutation. However, mechanisms for the emergence of other language concepts remain unclear.

According to Chomsky meaning are hard-wired already somewhere inside the computational device. In his opinion meaning are independent of social purposes, arising from genetically determined syntactic structure.

Chomsky's theory has been challenged by many authors through theory theory (Gopnik, 2003; Gopnik, 2012) and usage-based theory (Tomasello, 2009), for example.

According to Tomasello (2009), the innate grammatical module does not help a child assimilate his first language as much as different mechanisms of thinking, such as the ability to distribute objects into categories and to understand how they are related. Children begin by mastering simple grammatical schemes and then comprehend the rules that govern them. According to usage-based theory (Tomasello, 2009), children intuitively guess the intentions of other people. A child learns standard schemes by watching other people. Thus, in the author's opinion, the cognitive abilities of the brain serve as decisive factors of language learning. However, in light of that stated above with respect to problems of knowledge acquisition, these abilities must be explained. What, for example, does the term "assimilate" mean? How specifically does cognitive ability work with language when a child is exposed to a new environment?

According to theory theory (see for example, Gopnik, 2003; 2012), while obtaining knowledge about the world, children use the same tools as adults do when testing scientific principles. Children develop theories. According to the author (Gopnik 2003, 2012), infants are born with initial innate theories. However, the range of theories that can be generated is much wider than the limited set of representations that are possible from innate views. We learn by modifying, revising and eventually replacing earlier theories with later ones.

Theory theory is primarily applied to our everyday knowledge of the world, to our everyday understanding of biology, physics, psychology etc.

A significant drawback of the theory theory and the usual theory of language is their lack of basis in models and inaccurate definition of terms. What, for example, occurs when modifying one's original knowledge? How do we create new theories?

From this connection a question arises as to where theories of science come from (especially in the case of mathematics). Wigner (1967) explored the inexplicable effectiveness of mathematics. According to Wigner, mathematics is smarter than us. This means that the laws of mathematics and physics existed independently of humans before humans discovered them.

We show that the creation of concepts and generalizations is contradictory in principle regardless of the material basis of an intellectual system.

Creation of new concepts, abstraction and inductive logic

Let us consider several properties of thinking that are considered classical: the ability to generalize, intuition, abstraction, the creation of new concepts, and others. These concepts are at the core of human intellect and perhaps of the intelligence of animals. On the other hand, these properties are highly difficult (if not impossible) to apply on a computer (an intelligent system). Indeed, for example, abstraction modeling methods (Goertzel, 1993) inevitably suggest that a concept that should appear as a result of such an abstraction must be a priori defined in a mathematical model. Consequently, the implementation of such a model within an intellectual system also does not lead to the formation of a real abstraction.

In classical experiments with animals and young children (Gellerman, 1933) their abilities to form general concepts were measured. It is shown that children can identify the notion of "triangularity" in consideration triangles of different textures, colors, etc. That is, children (and some animals) are able to identify common geometric shapes and to recognize these commonalities in a new figure. Note, however, that when an image is recognized, it is translated into a mathematical (abstract) form. That is, along with other characteristics of an image, the property of "triangularity" is also recognized. Otherwise, it is impossible to explain how such a notion could have "formed" if it had not been programmed in advance. As is shown above, the trial and error method can be effective only when it is based on a hypothesis that is close to the truth a priori. Therefore, the trial and error method implicitly includes the Bayesian approach.

Lake and co-authors (2015) noted that two features of human thinking cannot be realized by computer systems. A human can learn a new concept with one attempt while a machine must make dozens or hundreds of attempts. Even children can make generalizations with one attempt. While learning, individuals can identify new options, functions and categories. Machines cannot do this. How do people learn from such a small amount of data? To address this problem, the authors proposed a Bayesian learning program that in their opinion supports the generalization and learning of new concepts. In this case, new programs can be developed from components of past ones.

However, in this case, the generalization and creation of new concepts is something quite different. If we consider signs registered by a system as images, then they can be processed in various ways, and from them it is possible to form new images, etc. However, this new image does not have semantics, such that one does not know what it displays. Accordingly, it reflects not the creation of a concept but the creation of a new sign. Operations of "averaging" signs can of course also be carried out. As a result, from a set of signs one can identify common features and form an "average" sign. However, as what it means is not known (the semantics are not averaged), the system cannot take any adequate action in this case. Thus, a generalization cannot be made at the syntactic level, as the semantics of these symbols are determined in a completely different way.

Consider for this connection the operation of "generalization." Let there be images in the external environment. After recognition, they correspond to concepts A_1, ... A_n of the internal language. When an image is not recognized, it is impossible to carry out operations related to semantics (i.e., it is discarded). Generalization means that all concepts belong to a certain set. In this case, two options are possible. Either they already belong to it and a priori there are operations adequate for this set (i.e., what to do with this set as the whole), or they do not belong to this set. In the latter case, the operation of replenishing the set with the object is not defined.

The problem of knowledge acquisition can also be considered based on logic. According to classical logic, to realize a proposition (that includes a subject, predicate, bundle and quantifier), all of its components must be a priori determined. This is often implicitly implied, but for our purposes it is necessary to note the importance of such an a priori definition. It is impossible to form a judgment of an unknown (unrecognized) object. Logical conclusions can be divided into inductive and deductive conclusions (see, for example, Baronett, 2008). It is assumed that deductive conclusions are strict while inductive conclusions are only probabilistic. Bertrand Russell noted (1998) that deduction does not give rise to new knowledge. New knowledge can only be obtained through induction. The problem of inductive inference has been raised repeatedly in

the field of philosophy. In particular, David Hume denied the validity of induction completely. On this connection Bertrand Russell remarked: "What these arguments prove is ... that induction is an independent logical principle that cannot be deduced from experience or from other logical principles, and that without this principle, science is impossible."

The analogy in logic is also regarded as a form of inference (see, for example, Baronett 2008), though its conclusions on induction are probabilistic. This probability value can be calculated from the Bayesian formula. However, as noted above, for this probability to be as great as possible, "good" hypotheses on which the Bayes formula itself is based must be applied. It remains unclear what are the mechanisms of a good hypothesis generation. The existence of such a hypothesis has only been postulated.

In the framework of formal (mathematical) logic, induction is also present (e.g., in Peano arithmetic). However, such induction can only work when all of its variants are determined in advance. Mathematical induction cannot work with objects that are not defined. Sometimes such a definition is given implicitly, but it does exist in any case.

Within the framework of logic, complete and incomplete induction are distinguished. Most often, induction is incomplete, and as shown above an unambiguous conclusion cannot be drawn from it.

In the case of complete induction we draw conclusions from a full enumeration of species of a certain genus on the whole genus; it is obvious that in using such a method of inference we obtain a quite reliable conclusion, and this method of inference cannot raise any doubts. In identifying the subject of a logical group with objects of particular judgments, we can apply this definition to the whole group.

The scheme of complete induction can be written as follows:

1. Set Ω consists of elements $\Omega_1, \Omega_2, ..., \Omega_k$.

2. Where Ω_1 has characteristic ψ, where Ω_2 has characteristic ψ, and where all elements of Ω_3 to Ω_k have characteristic ψ.

3. Consequently, all elements of set Ω have characteristic ψ.

In relation to the problem of knowledge acquisition, this situation can be interpreted as follows: in the case of acquiring knowledge, we always deal with a new object. As shown above, the reliability of inductive outputs is directly dependent on a priori information on this object. When a priori information is completely absent, no inductive conclusions can be drawn on the object.

If, from this point of view, we consider complete induction, it becomes equivalent to mathematical induction from which it is possible to draw definite conclusions, but induction can be complete only when all the elements of sets are determined a priori. In this case, no acquisition of knowledge (and understanding, the creation of new concepts, etc.) can occur in principle.

In general, we can conclude that mathematical logic and set theory are most closely related to thinking as such. The paradoxes of set theory and mathematical logic reflect the incompleteness of our knowledge. Namely, our thinking occurs partly in a latent form. At the same time, part of our thinking is not realized. Thus, some of the mathematical evidence is implicitly implied. We simply know that mathematics (e.g., an infinite set of real numbers) is constructed in this manner and not otherwise.

Consider the abstraction of identification as a means of forming general abstract concepts. This method involves taking into account only differences that for a given situation and for one reason or another are essential and in ignoring others and that are insignificant when considering initial objects. Initial objects that differ in an insignificant way are considered to be the same, and in terms of speech aspects, the abstraction of identification is manifested in the fact that two identical source objects when identified can be referred to as the same abstract object by applying the corresponding term. The following examples can be considered

- The identification of identical figures leads to the generation of an abstract figure (an abstract word);
- The identification of an equivalent fundamental sequence of rational numbers leads to the generation of a real number;
- The identification of isomorphic groups leads to the generation of an abstract group.

However, the word leads does not say anything about mechanisms facilitating the appearance of such concepts. It is much more logical to assume that concepts of "abstract letters," "abstract groups," etc.

already exist. As a special case, they must contain an isomorphic group, a rational number, and a letter. Therefore, when abstracting, a human only discovers such general concepts and does not create them.

Epistemology identifies various kinds of correct statements (see for example, Audi, 2011). There are self-evident statements and statements that are based on experience. Concerning the latter, however, one must bear in mind that all empirical must be recognized. Therefore, it is implicitly based on hypotheses that are postulated. That is, it is based on a priori knowledge. Analytical and synthetic correctness are also considered. Analytical correctness is obtained from proof. However, this is then just part of mathematical logic, for which the acquisition of new knowledge is impossible. Self-evident statements are innate by definition.

The problem of knowledge acquisition as a part of philosophy has a long history. In particular, the notion that knowledge is to some extent predetermined arose a very long time ago.

Socrates believed that soul is immortal (and repeatedly incarnated) and that the acquisition of knowledge is only an anamnesis. According to the ideas of Spinoza and Leibniz, thinking is a spiritual automaton that produces truth and that works according to a pre-set, pre-determined program. However, this raises questions as to where in such automaton errors come from. Descartes postulated that there is freedom of will according to which decisions are not predetermined.

Kant considered such categories innate and that the creative method works based on them. However, what is the creative method? This term remains poorly defined and does not clarify mechanisms of the emergence of new knowledge.

Kant asserted: "If we have a green leaf before us, our mind, comparing it with other leaves and considering through the prism of the category of universality, develops the concept of a leaf in general with the help of a productive imagination." This raises questions regarding what productive imagination is. What is universality? How does the mind define a category? Kant could not provide answers to these questions. To date progress in this area has been limited.

Any reasoning is formulated based on elementary operations of a certain language. Elementary logical operations can be performed by classical and quantum systems. In quantum mechanics, operations are performed on qubits. Simple unitary operations over qubits are called quantum logical gates. For example, when a qubit passes from one state to another, $|0\rangle \to |0\rangle$, $|1\rangle \to \exp(i\omega t)|1\rangle$ then it is said that after time t a qubit acts on the gate

Any reasoning is formulated on the basis of elementary operations within a certain language. Elementary logical operations can be performed by both classical and quantum systems. In quantum mechanics, operations are performed on qubits. Simple unitary operations over qubits are called quantum logical gates. For example, if a qubit passes from one state to another, $|0\rangle \to |0\rangle$, $|1\rangle \to \exp(i\omega t)|1\rangle$ then it is said that after a time t the gate

$$P(t) = \begin{pmatrix} 1 & 0 \\ 0 & \exp(i\omega t) \end{pmatrix}$$

acted on a qubit.

As other elementary quantum gates one can identify "equivalence," "not," etc. All of these gates act on one qubit and can be realized by means of some Hamiltonian of the Schrodinger equation. More detailed formalisms of quantum logic are described for example in (Haven and Khrennikov, 2017).

It is necessary, however, to emphasize that problems of justifying logic are not related to which particular model of logic (quantum or classical) is used. Thus, the problem of knowledge acquisition from a logical point of view cannot be solved within the framework of a given language either based on quantum theory or based on classical logic. Rather, for any language there is no means to justify the acquisition of knowledge.

Can neurophysiology provide an answer to the question of understanding and thinking?

If behavioral programs are innate, this in one way or another should be reflected in the structure of the brain.

The relationship between congenital programs of behavior and the structure of the brain is discussed in a previous book (Sverdlik, 2016). According to the author, the whole organism can be conditionally divided into two axes. The first axis includes our internal organs and tissues, and the second axis, the external axis, includes the skin and the sense organs. Accordingly, we can distinguish two divisions of the nervous system as follows: the autonomic (serving

the inner life of the organism) and the somatic (serving the external life) nervous system.

Different divisions of the brain are more or less involved in the process of the perception of the surrounding world and thinking.

The hypothalamus is the central regulator of the internal axis. The hypothalamus regulates the endocrine activity of the brain and the homeostasis of the body including processes such as food control, protective behavior, thirst, thermoregulation, memory, emotions, etc. (see for example, Swanson 2000; Canteras 2002).

The amygdala is involved in emotional and social processes (discrimination of faces, etc.) (Kheirbeck and Hen, 2011; Todorov and Engell, 2008).

The cortical mantle of the large hemispheres is the regulator of the outer axis. Conscious abstract thinking is carried out by stereotypes of the external axis. The resolving power of the autonomic nervous system is low and that of the somatic nervous system is high. Conscious thinking operates in working memory, in which there is a simultaneous retention of several pieces of information.

In the author's opinion, the neurophysiological meaning of abstraction is that, from the long and entangled neural networks short networks (that generalize their content) are formed. Here, however, an important question arises about the mechanisms of the formation of such networks. This mechanism should be a consequence of some physical processes.

According to the author (Sverdlik, 2016), the goal of abstraction is to find something common, which unified different information. The cortex works inaccurately, so the precision of mathematics does not depend on it. The limbic system is the basis of emotions. It is possible to attribute some neurobiological address to emotions. The emotional brain is the place in which the inner axis reaches the maximum contact with the external axis. There, the body and the cortex (algorithmic mind) find many points of contact. Emotions invisibly participate in solving any abstract problems.

In the author's opinion, to find the right solution for any problem means to recognize it as the right one. Recognition is controlled by the limbic (sensory) system. Neurons of the limbic system generate a sense of recognition when external information coincides with inner information. At the same time, the bodily component of cognition is important.

According to the author, true discovery is impossible without the use of non-algorithmic instruments. The body is not capable of constructing logical chains, and internal information is encoded somewhere in the molecules. Intuition is the emergence of a solution without the participation of working memory.

At the same time, however, the question arises as to why the body should be considered non-algorithmic. After all, the algorithm, in the most general sense, should describe all continuous processes.

On the other hand, non-algorithmicity (if it exists) simply means that this knowledge already exists.

To address the questions of understanding and consciousness, Koch and Tononi (2008) took a neurobiological point of view. Studying the activity of the brain with the help of a tomograph, the authors singled out the neuronal correlates of consciousness. The authors proposed an informational theory of consciousness, the main provisions of which are the following two statements:

1. Consciousness is a high informativity. Every private conscious state excludes a large number of other states.

2. Conscious information is unified.

If you present different images to a computer, the understanding of what is depicted or what looks wrong would allow a computer to pass a Turing test. However, according to the authors, the creation of a universal system that understands what is in front of it is impossible. However, progress in solving complex problems is attributed to the construction of computers that are organized similarly to the mammalian brain.

It remains unclear, however, how the proposed postulates help to understand how understanding arises in the event that a human meets an unknown image. How does the complexity of the brain and the large number of connections between neurons lead to understanding?

Quantum intelligence and quantum metalanguage

A number of authors associate the solution of a number of problems of the intellect with quantum

mechanics. These approaches can be divided into main motivations, i.e., reasons that, in the opinion of the authors, make the application of quantum mechanics to the thinking process promising. In my opinion, the most significant motivations are as follows.

One of the important areas of the application of quantum mechanics to cognitive sciences is decision making. This is an actively developing field that includes many researchers. The main motivation here is the decision, with the help of quantum mechanics, of some paradoxes of decision making. Such paradoxes, for example, include the Ellsberg paradox, which is a violation of the sure-thing principle and others. In the opinion of different authors, it is quantum mechanics that makes it possible to explain decision making in situations that the classical probability theory cannot explain. The most important property of quantum probabilities is that, in addition to the usual classical probabilities, they contain interference terms related to the wave character of quantum particles.

In the frame of «quantum-like consciousness» Khrennikov (2010a) proposed, that the brain performs "computation-thinking" by using algorithms of quantum computing in the complex Hilbert space of mental quantum-like models. Evolution of mental wave function in the frame of this model can be described with the Schrodinger equation:

$$i\frac{d\Psi}{dt}=\hat{H}\Psi.$$

Here

$\hat{H}=\frac{\hat{b}^2}{2}+V(\hat{a})$ - is the operator of «mental energy», $\hat{a}$ and $\hat{b}$ - are two self-adjoint operators, and $V(\hat{a})$ - is the «mental potential». As will be shown below, the Hilbert space can be represented in the form of a subspace in which the metalanguage operates.

The application of Bohmian version of quantum mechanics to the processes of thinking is caused by the fact, that the pilot wave (on the concept of which Bohm's quantum mechanics is largely based) can, to some extent, be considered the simplest form of consciousness. For example, Pylkkanen (2016) believes that some properties of quantum systems can help to understand the properties of consciousness. The pilot wave carries active information and, in a sense, informs the particle about its surroundings. According to the author, consciousness exists only because there is a higher level of information. This last remark can be attributed not only to consciousness but also to thinking in general. In the works of Khrennikov, mental information fields were defined in the "information space", reflecting a hierarchic tree-like representation of information by cognitive systems (mathematically, such mental trees can be represented with p-adic numbers (Khrennikov, 2010b).

A second quantization method is also promising for quantum-like modeling in which particle creation and annihilation operators are used (Bagarello, 2012; Bagarello *et al*., 2017). In this case, these operators can be used to create or destroy the simplest information units of decision making, such as "Yes" or "No".

The application of quantum mechanics to thinking can also be justified by the fact that the brain itself, on the basis of which thought is realized, functions as a quantum system. In previous works (Melkikh 2014b; Melkikh and Meijer, 2018), it was shown that the work of neurons at the molecular level is contradictory. This phenomenon is reflected in the fact that the accuracy of the biochemical reactions of protein-ligand and protein-protein interactions, as well as the accuracy of the folding of proteins and DNA (RNA), cannot be justified by the presence of short-range potentials between biologically important molecules. In this case, the entangled (in the classical sense) and the inoperable states of macromolecules should be realized with an overwhelming probability. In particular, this will lead to the impossibility of the efficient transport of substances both inside neurons and through their membranes. This phenomenon (formulated by the authors as a generalized Levinthal's paradox) requires a fundamental revision in the understanding of the mechanisms of the work of cells at the molecular level. To solve the paradox, the authors proposed a quantum model of intermolecular interactions. The most significant point of this model is the long-range interaction, which ensures the efficient operation of intracellular molecular machines. Here, the motivation for using quantum mechanics is based on the fact that classical mechanics cannot, in principle, provide such an interaction.

As noted above, the motivation of Penrose's proposals on the application of quantum mechanics is that he connects mental processes with the collapse of the wave function, which, in the author's opinion,

is caused by the effects of gravity. Moreover, non-computability plays an important role. According to the author, the ability to understand cannot be formalized with the help of a particular set of rules. The author comes to the conclusion that "for the establishment of mathematical truth, mathematicians do not apply knowingly justified algorithms."

Is it possible to conclude from this idea that understanding in principle is not formalizable? Such a conclusion cannot be made, since unprovability occurs only within a certain language. There is no reason to say that the language we use is closed and unchanged.

Hence, what follows is a natural logical conclusion that the concept of truth (adequacy in the broad sense) in relation to certain constructions is innate. This concept is just as axiomatic as the rules of the language within which there is a discussion (modeling) of truth.

As for quantum gravity, its necessity for the operation of microtubules in neurons can hardly be justified, since the effects of quantum gravity should appear at much higher energies. Nevertheless, the concept of microtubules and the cytoskeleton as quantum systems is currently developing (see for example, Cocchi *et al*., 2017).

The ideas of Zizzi can, in a sense, be considered a continuation of Penrose's ideas. According to Zizzi (2012), there are three ways of expressing mental activity: classical computer, quantum computer and a non-computational way. In the author's opinion, we do not realize quantum calculations, but classical ones are conscious. The third way is connected with meta-knowledge - intuition, which is controlled by the quantum metalanguage. It is this language that distinguishes people from computers. People, in addition to logical rules, have meta-logical rules. We can agree with this, but it is necessary to clarify that these rules do not fundamentally differ from logical rules; they are simply hidden and manifest only in certain situations. Such properties of metalanguage will be discussed below in more details.

According to the author (Zizzi, 2012), a mathematician can uncomputably assert what Gödel's theorem considers unprovable. We note that Gödel's theorem addresses a closed logical system into which the hidden metalanguage is not included. Within the framework of the metalanguage, these statements could be completely provable. Although there certainly will be other unprovable statements, this is another matter.

In the author's opinion, the quantum metalanguage governs our own language. When a mathematician claims the correctness of some statement G(F), in reality, he operates at the level of the quantum metalanguage.

As will be shown in the next section, the concept of quantum metalanguage is very important and can be used to solve the problem of knowledge acquisition and to create a new paradigm of thinking in general.

Discussion

Quantum metalanguage and innate programs of behavior

Discussion the problems of knowledge acquisition in the context of various sciences (epistemology, philosophy of science, neurocomputing and others) leads to the need for a new cognitive synthesis. This synthesis must be connected with a new answer to the following basic question of cognitive sciences: where does the knowledge come from? To this end, it is necessary to take advantage of all that has been achieved in various cognitive sciences. Previously, for example, Griffits (2015) proposed to consider the active use of computers and calculations in the cognitive sciences, which occurs at present as a new cognitive revolution. However, such use alone does not provide new answers to fundamental questions but only helps to speed up the experiments. The analysis of artificial intelligence, conducted in Section 5, shows that the computer does not provide new knowledge.

To more accurately define the concepts related to thinking, the adequacy of human behavior has been suggested (Melkikh, 2014a). Indeed, we can verify only experimentally whether a human has realized something new and whether he could create a new abstract concept. In this case, the only way to verify this is to determine how the behavior of a human (an arbitrary intellectual system) has changed as a result of understanding. If this behavior has become, to some extent, more adequate than before understanding, then we can say that understanding really occurred. If no adequacy is observed, then there is no way to verify that understanding really occurred.

If we consider an intellectual system in the form of an automaton with given properties, then the adequacy of its behavior can be estimated, for

example, by the gain that this system obtains from interaction with the environment. The following model of behavior is proposed (Melkikh, 2014a):

$$\varphi(t+1)=\Phi(\varphi(t),s(t+1))$$

$$f(t)=F(\varphi(t)).$$

Where, φ describes internal states of automaton, f is the action of automaton, t is the discrete time and s is the input variable.

The definition of knowledge acquisition of in accordance with the proposed model is follows:

Definition. This will be understood by the acquisition of knowledge of such a process, followed by the emergence of an environmentally unrecognized image c* not including preparation of the systems and only to provide the system with information regarding the external environment through receptors, which results in a behavior of the system that is estimated to increase the expectation of winning at some preset value of W_0:

$$W(c^*)>W_0.$$

The assignment of the image to unrecognized objects indicates that there is no prior information about the object.

The hypothesis of the congenital behavior of animals is proposed (Melkikh, 2014), the main provisions of which are the following:

1. Innate programs of behavior are formed on the basis of genes during ontogenesis (morphogenesis).

The formation process of these programs is not obvious because, for example, in an organism, there is a clear discrepancy between the information contained in the genes and the complexity of the nervous system. However, the modeling of morphogenesis is a separate problem and is not considered in this paper.

2. As a result of receptors receiving signals from the environment, remembering and recognition occur.

Memory of the time or spatial distribution of the signals used by the organism to run a priori behavior programs is the most adequate for the observed situation. If recognition is not possible (the appeared image is new), then either no a priori programs start or the programs that correspond most closely (on the tree of recognition) to the recognized image are run.

3. At the presence of uncertainty in the environment or errors in receptors aprioristic programs run only after several repetitions.

Repetition can reduce random error in the registration signal. This situation is largely similar to the decrease of random errors in measurements.

4. Another response to the uncertainty of the environment is a trial and error method, which allows the system to select the most appropriate aprioristic behavior program.
5. An interaction with another organism that transfers the experience of one organism with the environment (see items 2-4) to another is only possible if both the organisms have the same a priori programs relating to the present situation.
6. All programs of animal behavior are innate. For technical systems, they are aprioristic programs.

Thus, animal learning caused by uncertainty in the environment is the method of choice for the a priori (innate) program, which is most appropriate for a certain situation. The new behavior program cannot be formed on the basis of an experience or interaction with other animals. This hypothesis may well be extended to the process of human cognition.

Thus, to learn is to choose one of the a priori programs most appropriate to a given case. To quickly understand the situation (insight) means to immediately choose such a program. To learn by trial and error means to not immediately find such a program because of interference, environmental uncertainty, etc.

To solve the problem of knowledge acquisition, consider the concept of metalanguage introduced by Tarski. It is a higher-level language that works with other languages. Some operations can be used to generate new languages from given languages. At the same time, operations such as concatenation, intersection, association and others are defined.

Suppose that such operations are defined on our language. These operations are defined in some higher-level language L(Fig. 4).

Language L_2 (our) is constantly replenished due to the innate quantum language L_1 according to the rules of language L . Thus, we can build the following logical chain, modeling thinking as follows:

Changes in the environment → registration by the receptors → recognition with the L_2 language

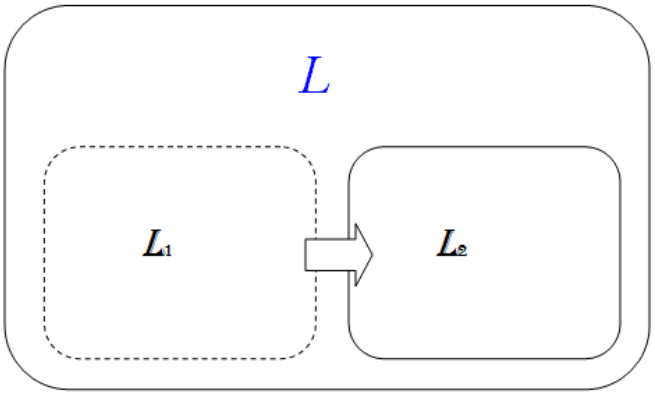

Figure 4. Higher-level language

→ if not → recognition with the metalanguage → operations on the L_2 language involving the image, such as generalization and the formation of new concepts → the introduction of a new character symbol in the L_2 language (together with possible operations) → new knowledge. This sequence of actions can also be presented in the form of a flowchart (Fig. 5):

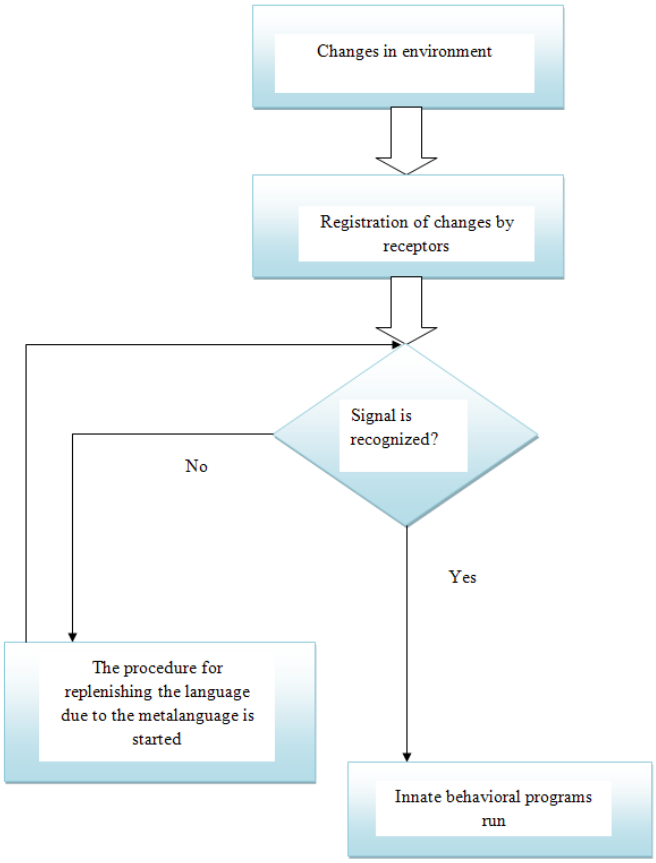

Figure 5. Logical chain, modeling thinking

Quantum mechanics represents the natural conditions for the formulation of the metalanguage, the discreteness of energy and states. Within the framework of quantum mechanics of particles, the following set acts as a language:

$$\left\{\Psi, \hat{H}\right\}.$$

That is, the Hamiltonian can be considered as a set of operations on wave functions. However, generally speaking, to construct a quantum model of an arbitrary system, it is necessary to consider the quantization of the fields.

The metalanguage operates in a more general space, part of which is the Hilbert space of wave functions. The elements of the metalanguage can also be the creation and annihilation operators applied to various elements of the language. With the help of such operators in the language, new concepts are selected from a list of a priori existing concepts.

For effective functioning, a quantum metalanguage must have a large information capacity in an orders of magnitude greater than the capacity of any of the known languages. At first glance, such a requirement contradicts the comparatively small informational capacity of genes. Indeed, genomes of higher organisms contain only about a gigabyte of information, which is clearly insufficient to manage such complex behavior as human behavior. This contradiction can be solved on the basis of the assumption that the interaction between biologically important molecules (including neurons) is essentially a quantum effect (Melkikh, 2014b; Melkikh and Meijer, 2018). As noted above, the essence of the proposed model is that long-range potentials act between biologically important molecules, which allow the modulation of their movement. This additional interaction allows us to increase the possible information capacity of molecules by orders of magnitude, as well as cells in general, which is not related directly to genes. Experiments were proposed to test this hypothesis in previous works (Melkikh, 2014b; Melkikh and Meijer, 2018).

Within the framework of the new concept of innate behavior, we can answer the following Kant's questions:

— How are synthetic judgments possible a priori?

— How is human freedom possible?

The first question can be answered as follows: a priori judgments are the consequences of the innate quantum metalanguage, which is adequate a priori to the existing reality.

The second question can be answered in such a way that what is considered free will is the result of unconscious metalanguage algorithms. If the basis of thinking is quantum mechanics, then this naturally explains the existence of probabilistic judgments (which seem to be arbitrariness and freedom). As is known, the outcome of an experiment with quantum particles can only be predicted probabilistically. This property, with respect to operations over elements of the language, can be treated as freedom of will.

Consider how a human learns a new language within the framework of the concept of innate behavior. From this point of view, the presence of Chomsky's proposed Language Acquisition Device is quite natural. This is nothing other than an innate program of behavior. Moreover, it can be said that this is not a mechanism for extracting the language but rather a mechanism for including an innate program that implements the language. This applies, of course, not only to the native language of any person but to any language in general. However, we know that the ability to learn languages is different among all people. The mechanism of this phenomenon within the framework of the metalanguage can be thought of as the following: all people have the same congenital programs, however, they start differently in the development process. This process is largely determined by the environment, which leads to the fact that some programs are not run at all. In this sense, the programs under which we operate are operational programs. Most of the programs are hidden and not operational. They cannot manifest themselves in the course of a human life.

Let the child in the appropriate situations (surrounding objects) be given words that correspond to these subjects. After numerous repetitions, the innate programs for working with these words are run. At the same time, the understanding of words is present a priori. The further generalization of concepts (words) occurs in a similar way; there is recognition of the fact that a given set of words refers to a particular (a priori) set that is characterized by some generalized concept. After recognition, the conclusion is drawn that the concept is formed. All programs of work with such an abstract concept exist a priori (they are innate).

Thus, if one considers Chomsky's approach to language acquisition from this point of view, then it can be called more correct than, for example, the cognitive approach or the theory theory, since the latter contains many vague terms, and the mechanisms for mastering new terms, understanding new concepts and so on are contradictory. However, Chomsky's approach also needs further development. If only a part of the language is innate, it is not clear where the remaining parts originate. In this case, it is logically consistent that the whole language is innate. This does not mean that the language in this case acts as a kind of rigid structure. A natural property of algorithms and computer programs written on their basis are operators of the following type:

if A then B, else

if C then D, else

if E then F else….

Thus, completely innate behavior allows to flexibly react to changes in the environment, including forming (which means choosing from ready-made concepts) various new concepts in different life situations.

Let us consider a simple example. How can a child create the concept of a square as a result of communication with the outside world? Let the child be surrounded by the following objects (Fig. 6):

At some initial stages of its development, generalization simply does not occur, and the child perceives these objects as different. However, in the future, the following occurs: the recognition of all these objects includes the congenital concept of a "square". That is, with their recognition, the child,

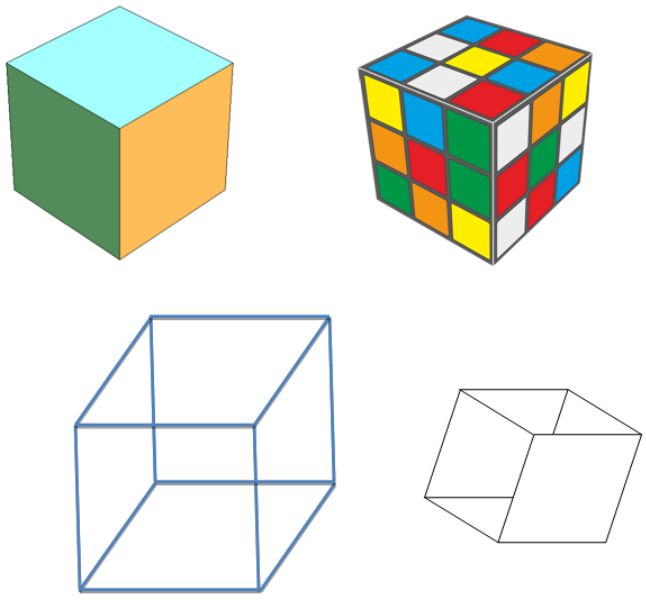

Figure 6. Different objects and concept of a square

in addition to other qualities of objects, recognizes the objects as belonging to a set of "squares". This set, with all its properties, is also innate, and the information and programs are written on an innate quantum metalanguage. At a certain stage (when a real need arises), these congenital programs are translated from the quantum metalanguage into the child's language. The concept of "square" thus begins to work, i.e. the child begins to use this concept in life and in the future to more adequately address emerging problems.

A number of experiments to test the innate behavioral programs were proposed earlier (Melkikh, 2014a; 2014b; Melkikh and Khrennikov, 2015). Planning and conducting such experiments is closely related to a precise understanding of what is innate (see, section 4).

In the end, only such experiments (or the like) can finally resolve the question of whether cognitive effects are quantum or quantum-like.

Thus, the new paradigm of cognitive sciences can be formulated in the form of the following provisions:

1. All human behavior is innate. It seems to us that, as a result of learning and gaining experience, we can create new concepts. However, we are just beginning to apply the concepts that were already available.

2. Bayesian learning is entirely based on good hypotheses. These hypotheses are good because they are part of innate programs of behavior that are adequate a priori to the surrounding world.

3. When we say that we understand something, it means that it seems new, recognized, and it corresponds to some innate standard.

4. Congenital behavior programs allow behavior to be flexible through multiple types of structures of "if ... then ...".

Conclusion

A systems analysis of the processes of understanding, generalizing and knowledge acquisition leads to the conclusion that these processes are contradictory. However, their detailed consideration is fundamentally important for understanding intelligence as such. The problem of the Chinese Searle room is considered, and it is shown that this problem can be naturally solved on the basis of the innate knowledge of the language. In this sense (meaning precisely the acquisition of knowledge, not feelings), a human is no different from a computer.

To solve these problems and paradoxes, a quantum metalanguage model is proposed that provides cognitive function only at the expense of innate behavioral programs. As a result, it is possible to explain consistently the mechanisms of the emergence of new concepts and generalizations. A priori correct judgments are possible because they are correct innately. As a result of training, only a choice of appropriate behavioral programs occurs, in which all new concepts, generalizations, and so on are already included a priori. The fact that artificial intelligence systems work successfully is due to only one thing - a significantly increased speed of pattern recognition. This speed now allows the solving of many tasks online. However, this does not mean that such intellectual systems are able to acquire knowledge or understand something new.

Appendix 1

Connection between information, obtained in experiment and the number of measurements

In a previous paper (Melkikh, 2014a), the trial and error method as a series of experiments was considered, which will result in finding an adequate algorithm or the absence of an adequate algorithm. In information theory, the degree of uncertainty of experiment β with possible outcomes and their probabilities $p_1, p_2 \ldots p_k$ is usually characterized by the Shannon entropy:

$$H(\beta) = -\sum_{i=1}^{k} p_i \log_2 p_i .$$

If this value equals zero, the outcome of the experiment is known in advance. Any measurement or supervision α prior to experiment β can limit the quantity of possible experimental outcomes and thus reduce a degree of its uncertainty. The fact that the realization of experiment α reduces a degree of uncertainty in experiment β is reflected through the conditional entropy H(β/α) of experiment β under the condition of α being less than (more precisely, less than or equal to) the initial entropy of the same experiment. Thus, if experiment β does not depend on α, the realization of α does not reduce the entropy of β:

$$H(\beta/\alpha) = H(\beta),$$

If the result of α completely predetermines an outcome of β, the entropy decreases to zero: $H(\beta/\alpha) = 0$. Thus, the difference

$$I(\alpha,\beta)=H(\beta)-H(\beta/\alpha)$$

specifies how much the realization of experiment α reduces the uncertainty of experiment β. This difference represents the quantity of information in experiment β, which is contained in experiment α. Here, information (according to the information theory) is understood as the uncertainty that is removed when a message is received.

There is an interrelation between the number of experiments and the information received as a result of each experiment. For example, if the experience is found with an adequate algorithm, each experience corresponds to receiving one bit of information. Suppose that the system has n degrees of freedom, each of which may be in one of m states. If, at the beginning of the experiment, there is no a priori information to find an adequate algorithm, then we obtain the average number of steps required to achieve it, which is equal to the total number of possible states:

$$N=m^n\,.$$

If n is sufficiently large, then the number of steps is exponentially large, even for m = 2.

The number of steps corresponds to Shannon entropy:

$$H=\log_2 N=n\log_2 m\,.$$

In this case, each bit of a priori information reduces the search area in half.

References

Abu-Mostafa YS, Magdon-Ismail M, Lin H-T. Learning from data. Pasadena: AMLbook.com, 2012.

Aerts D, Sozzo S, Veloz T. Quantum structure of negation. Frontiers in Psychology 2013; 6: 1447.

Aerts D, Gabora L, Sozzo S, Veloz T. Quantum structure in cognition: fundamentals and applications. arXiv:1104.3344v1, 2011.

Agrawal AA. Phenotypic Plasticity in the Interactions and Evolution of Species. Science 2001; 294: 321-326.

Audi R. Epistemology. Third edition. Routledge. Taylor and Francis Group. New York, London, 2011.

Bagarello F. Quantum dynamics for classical systems: with applications of the number operator. New York: J. Wiley, 2012.

Bagarello F, Basieva I, Khrennikov A. Quantum field inspired model of decision making: Asymptotic stabilization of belief state via interaction with surrounding mental environment. Journal of Mathematical Psychology 2017; 82: 159-168.

de Barros AJ. Quantum-like model of behavioral response computation using neural oscillators. Biosystems 2012; 110: 171-182.

Baronett S. Logic. Upper Saddle River, NJ: Pearson Prentice Hall, 2008.

Barsalou LW. Perceptions of perceptual symbols. Behavioral and Brain Sciences 1999; 22(4): 637-660.

Basieva I, Pothos E, Trueblood J, Khrennikov A, Busemeyer J. Quantum probability updating from zero prior (by-passing Cromwell's rule). Journal of Mathematical Psychology 2017; 77: 58-69.

Basti G, Capolupo A, Vitiello G. Quantum field theory and coalgebraic logic in theoretical computer science. Progress in Biophysics and Molecular Biology 2017; 130(A): 39-52.

Bishop JM, Nasuto SJ, Coecke B. 'Quantum Linguistics' and Searle's Chinese Room Argument. Müller V. (eds) In "Philosophy and Theory of Artificial Intelligence". Studies in Applied Philosophy, Epistemology and Rational Ethics, vol 5. Springer, Berlin, Heidelberg, 2013: 17-28.

Bonawitz E, Denison S, Griffiths TL, Gopnik A. Probabilistic models, learning algorithms, and response variability: sampling in cognitive development. Trends in Cognitive Sciences 2014; 18(10): 497-500.

Butcher LM, Davis OSP, Craig IW, Plomin R. Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500K single nucleotide polymorphism microarrays. Genes, Brain and Behavior 2008; 7(4): 435-446.

Canteras NS. The medial hypothalamic defensive system: Hodological organization and functional implications. Pharmacology, Biochemistry and Behavior 2002; 71: 481–491.

Chalmers D. The conscious mind. New York. Oxford University Press, 1996.

Chomsky N. Syntactic Structures. The Hague/Paris: Mouton, 1957.

Chomsky N. Aspects of the Theory of Syntax. MIT Press, 1965.

Chomsky N. The minimalist program. The MIT Press. Cambridge, Massachusetts, 2015.

Cocchi M, Minuto C, Tonello L, Gabrielli F, Bernroider G, Tuszynski JA, Cappello F, Rasenick M. Linoleic acid: Is this the key that unlocks the quantum brain? Insights linking broken symmetries in molecular biology, mood disorders and personalistic emergentism. BMC Neuroscience 2017; 18: 38-48.

Coleman JRI, Bryois J, Gaspar HA, Jansen PR. et al. Biological annotation of genetic loci associated with intelligence in a meta-analysis of 87,740 individuals. Molecular psychiatry 2018; 8: 1.

d'Avila Garcez AS, Lamb LC, Gabbay DM. Neural-Symbolic Cognitive Reasoning, Cognitive Technologies. Springer-Verlag, Berlin Heidelberg, 2009.

Dennett D. Sweet dreams. Philosophical obstracles to a science of consciousness. A Bradford Book, The MIT Press, Cambridge, Massachusetts, and London, England, 2005.

Fodor JA. The Modularity of Mind: an Essay of Faculty Psychology. MIT Press, 1983.

Fodor JA and Pylyshyn ZW. Minds without meanings. An essay on the content of concepts. The MIT Press, 2016.

Garey M and Johnson D. Computers and Intractability: A Guide to the Theory of NP-Completeness. Freeman, San Francisco, 1979.

Gellerman LW. Form Discrimination in Chimpanzees and Two-Year-Old Children: I. Form (Triangularity) Per Se, The Pedagogical Seminary and Journal of Genetic Psychology 1933; 42(1): 3-27.

Glanzman DL. Common Mechanisms of Synaptic Plasticity in Vertebrates and Invertebrates. Minireview. Curr Biol 2010; 20: 31-36.

Goertzel B. The structure of intelligence. A new mathematical model of mind. Springer-Verlag. New York, 1993.

Gopnik A. (2003) The theory theory as an alternative to the innateness hypothesis. Book chapter in: In L. Antony and N. Hornstein (Eds.), Chomsky and his critics. Oxford: Blackwells. Retrieved 2013-04-26.

Gopnik A. Reconstructing constructivism: Causal models, Bayesian learning mechanisms, and the theory theory. American Psychological Association 2012; 138: 1085–1108.

Gorlich D, Artmann S, Dittrich P. Cells as semantic systems. Bioch Bioph Acta 2011; 1810: 914–923.

Gottfredson LS. Mainstream Science on Intelligence (editorial). Intelligence 1997; 24: 13–23.

Griffiths TL. Manifesto for a new (computational) cognitive revolution. Cognition 2015; 135: 21-23.

Hameroff S and Penrose R. Consciousness in the universe: A review of the 'Orch OR' theory. Phys Life Rev 2014; 11: 39-78.

Haven E and Khrennikov A. The Palgrave handbook of quantum models in social science. Applications and grand challenges. Palgrave Macmillan, London, 2017.

Hernandez-Orallo J. The measure of all minds. Evaluation of natural and artificial intelligence. Cambridge University Press, 2016.

Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh AM, Gale CR, Davies G, Deary IJ. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Molecular psychiatry 2018.

Kamsu-Foguem B, Tchuenté-Foguem G, Allart L, Zennir Y, Vilhelm Y, Mehdaoui H, Zitouni D, Hubert H, Lemdani M, Ravaux P. User-centered visual analysis using a hybrid reasoning architecture for intensive care units. Decision Support Systems 2012; 54(1): 496–509.

Kant I. Critique of pure reason. Cambridge University Press, 1998.

Khrennikov A. Classical and quantum mechanics on information spaces with applications to cognitive, psychological, social and anomalous phenomena. Foundations of Physics 1999; 29(7): 1065-1098.

Khrennikov A. Ubiquitous quantum structure: from psychology to finances, Springer, Berlin-Heidelberg-New York, 2010a.

Khrennikov A. Modelling of psychological behavior on the basis of ultrametric mental space: Encoding of categories by balls. P-Adic Numbers, Ultrametric Analysis, and Applications 2010b; 2, N. 1: 1-20.

Khrennikov A. Quantum-like model of processing of information in the brain based on classical electromagnetic field. Biosystems 2011; 105(3): 250-262.

Kheirbeck MA and Hen R. Dorsal vs ventral hippocampal neurogenesis: implications for cognition and mood. Neuropsychopharmacology 2011; 36(1): 373–374.

Koch C and Tononi G. Consciousness as integrated information. Biological Bulletin 2008; 215(3): 216-242.

Korf J. Quantum and multidimensional explanations in a neurobiological context of mind. The Neurocsientist 2015; 21(4): 345-355.

Lake BM, Salakhutdinov R, Tenenbaum JB. Human-level concept learning through probabilistic program induction. Science 2015; 350: 1332-1338.

LeCun Y, BengioY, Hinton G. Deep learning. Nature 2015; 521: 436–444.

Legg S, Hutter M. (2007) A collection of definitions of intelligence. Advances of Artificial General Intelligence: Concepts, Architectures and Algorithms. Eds. Goertzel, B., Wang, P., IOP Press: 17-24.

Lorenz KZ. (1950) The comparative method in studying innate behavior patterns. in "Physiological Mechanisms of Animal Behavior". 221-268. Cambridge University Press.

Luger GF. Cognitive Science: The Science of Intelligent Systems. Academic Press, San Diego and New York, 1994.

Luger GF. Artificial intelligence. Structures and strategies for complex problem solving. Fourth edition. Addison Wesley, 2003.

Maruyama Y. AI, Quantum Information, and External Semantic Realism: Searle's Observer-Relativity and Chinese Room, Revisited. InFundamental Issues of Artificial Intelligence 2016; pp: 115-127. Springer, Cham.

Mathôt S, Melmi J-B, Van Der Linden L, Van Der Stigchel S. The mind-writing pupil: a human-computer interface based on decoding of covert attention through pupillometry. PLoS ONE 2016; 11 (2): e0148805.

Melkikh AV. The No Free Lunch Theorem and hypothesis of instinctive animal behavior. Artificial Intelligence Research 2014a; 3(4): 43-63.

Melkikh AV. Congenital programs of the behavior and nontrivial quantum effects in the neurons work. Biosystems 2014b; 119: 10-19.

Melkikh AV. Quantum information and the problem of mechanisms of biological evolution. BioSystems 2014c; 115: 33-45.

Melkikh AV and Meijer DKF. On a generalized Levinthal's paradox: the role of long- and short range interactions in complex bio-molecular reactions, including protein and DNA folding. Progress in Biophysics and Molecular Biology 2018; 132: 57-79.

Melkikh AV and Khrennikov A. Nontrivial quantum and quantum-like effects in biosystems: unsolved questions and paradoxes. Progress in Biophysics and Molecular Biology 2015; 119(2): 137-161.

Melkikh AV and Mahecha DS. On the Broader Sense of Life and Evolution: Its Mechanisms, Origin and Probability across the Universe: Journal of Astrobiology & Outreach 2017; 5(3).

Miller GA. The cognitive revolution: a historical perspective. Trends in Cognitive Sciences 2003; 7(3): 141-144.

Minsky ML. The Society of Mind. William Heinemann Ltd, London, 1987.

Noble D. The Music of Life. Biology Beyond Genes, Oxford University Press, Oxford, 2006.

Osherson D, Stob M, Weinstein S. Systems that learn. MIT Press. Cambridge MA, 1986.

Pask G. The cybernetics of human learning and performance. Hutchinson, 1975.

Penrose R. Shadows of mind. A search of the missing science of consciousness. Oxford University Press. New York, Oxford, 1994.

Pothos EM and Busemeyer, J.M. Can quantum probability provide a new direction for cognitive modeling? Behavioral and brain sciences 2013; 36: 255–327.

Pylkkanen P. Can Bohmian quantum information help us to understand consciousness? In: Atmanspacher H., Filk T., Pothos E. (eds) Quantum Interaction. Lecture Notes in Computer Science, Springer, Cham 2016; l 9535: 76-87.

Russell B. History of western philosophy. Routledge. Tailor and Francis Group. London and New York, 2009.

Samuels R. Innateness in cognitive science. Trends in Cognitive Sciences 2004; 8(3): 136-141.

Searle JR. Minds, Brains, and Programs. Behavioral and Brain Sciences 1980; 3(3): 417–457.

Silver D, Hubert T, Schrittwieser J, Antonoglou I, Lai M, Guez A, Lanctot M, Sifre L, Kumaran D, Graepel T, Lillicrap T, Simonyan K, Hassabis D. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. arXiv:1712.01815v1. 2017.

Smith CU. The 'hard problem' and the quantum physicists. Part 1: the first generation. Brain Cogn 2006; 61: 181–188.

Smith CU. The 'hard problem' and the quantum physicists. Part 2: modern times. Brain Cogn 2009; 71: 54–63.

Sverdlik A. How our emotions and bodies are vital for abstract thought: perfect mathematics for imperfect minds. Taylor and Francis. New York, 2018.

Swanson LW. Cerebral Hemisphere Regulation of Motivated Behavior. Brain Research 2000; 886: 113–164.

Sze V, Chen Y-H, Yang T-J, Emer J. Efficient processing of deep neural networks: a tutorial and survey. arXiv:1703.09039v2. 2017.

Tomasello M. Constructing a language. A usage-based theory of language acquisition. Harvard University Press. Cambridge, Massachusetts, 2009.

Todorov A and Engell AD. The role of the amygdala in implicit evaluation of emotionally neutral faces*. Social Cognitive and Affective Neuroscience 2008; 3(4): 303–312.*

Vilhelm C, Ravaux P, Calvelo D, Jaborska A, Chambrin M-C, Boniface M. Think!: a unified numerical-symbolic knowledge representation scheme and reasoning system. Artificial Intelligence 2000; 116(1–2): 67–85.

Wigner E. The unreasonable effectiveness of mathematics in natural sciences. In Symmetries and Reflections. Indiana University Press, 1967; pp: 222-237.

Whitman DW and Agrawal A. What is phenotypic plasticity and why is it important? Phenotypic Plasticity of Insects: Mechanisms and Consequences (ed. by D. W. Whitman and T. N. Ananthakrishnan), Science Publishers, Enfield, New Hampshire 2009; pp: 1–63.

Yampolsky RV. On the Limits of Recursively Self-Improving AGI. In: Bieger J., Goertzel B., Potapov A. (eds) Artificial General Intelligence. AGI 2015. Lecture Notes in Computer Science, Springer, Cham, 2015; 9205.

Yampolskiy RV. Form seed AI to technological singularity via recursively self-improving software. arXiv: 1502.06512, 2015.

Yampolskiy RV. Detecting qualia in natural and artificial agents. arXiv:1712.04020, 2017.

Zizzi P. Non-algorithmic side of the mind. arXiv:1205.1820, 2012.