EPISTEMOLOGICAL, LOGICAL AND SEMANTIC PROBLEMS OF CREATING ARTIFICIAL INTELLIGENCE

by A.V. Sukhoverkhov and N.E. Spasova

From ancient times to this day, philosophers and scientists have sought to understand human thinking and consciousness. Modern science tries to model the human brain and create technical devices that copy mental processes. Of course, this is a very difficult task, since many aspects of brain structure and function are unknown. However, what we do know allows us to begin the task. In this article we discuss the possibilities of creating artificial intelligence (AI) and the theoretical and practical difficulties that arise in its development.

Scientists who develop AI proceed on the view that the mind is part of the natural world, best regarded as the mental workings of an embodied brain, subject to laws and rules that can be comprehended and thereby reproduced artificially. The notion that these rules are those of logic underpinned the contributions of Frege, Russell, Whitehead, Boule, Turing, Schenk, Newell, Simon and others [2].

In the latter half of the 20th century, digital computers and neural networks were developed. The first manipulated symbols according to preset programs (data processing), the second learned from its inputs (connectionism). Each had its advocates as the best way to model cognition.

Championing the computer were American scientists Newell and Simon who came to the conclusion that the sequence of bits of computers can represent not only numbers but also objects and features of objective reality, and semantic and subjective values. Dreyfus and Dreyfus, documenting the development of AI, wrote that Newell and Simon "put forward the hypothesis that the human brain and digital computer, being absolutely different in structure and mechanism, have at a certain level of abstraction one and the same functional description" [4, p. 402]. This led to the conclusion that both the living brain and the computer generate intelligent behavior. So arose the hypothesis about the "system of physical symbols" being a necessary and sufficient condition for the creation (modeling) of reasonable actions. The task was to develop simple symbol strings that denote the facts of the world and find the right logical connections between them. As a result, Newell and Simon managed to create a computer that could solve simple puzzles and prove small theorems. In their opinion, they managed to create technical devices that copy human learning and heuristics.

Advocates of connectionism, Hebb and Rosenblatt, sought to create a model of the human brain. Scientific support for neural networks came from neuroscience and mathematics. Neural networks consisted of interconnected artificial neurons in which connection strengths determining output were repeatedly altered by the input. These scientists wanted to create a technical device that could not only use its abilities, but also develop them. Rosenblatt called his equipment a perceptron. In it patterns were both stored and modified by experience. Such a device seemed to have acquired creative abilities. Luger writes: "Neural networks are trained not by adding new information to their knowledge base, but by modifying their general structure in response to information received from outside" [6, p. 372]. A feature of neural circuits is the ability to generalize, so the task of AI creators was to teach machines to correctly apply a pattern of data in new situations. An important feature of rational behavior of a living organism is its ability to learn; therefore, in order to bring the network closer to the human brain, it is necessary to develop in it the ability to learn. Rosenblatt set this very goal, but it turned out to be very difficult in practice. If the machine is given a second attempt to solve the problem, unlike a human, it will not remember its first solution and will carry out all the computational processes anew. Scientists have noticed one important difference between AI and human intelligence: in a person, more information makes a solution easier, while in the computer, the more information it contains, the more difficult it is to cope with tasks. Since the perceptron scientists, whose computers were overloaded with information, had to make cumbersome calculations even for the solution of the simplest tasks, their project, in the end, was not very fruitful.

Work on AI continued and approached the stage when it became possible to compare the work of the human brain with the computer. The most interesting way of comparison was the Turing Test. The computer and the person are asked the same questions, and the researcher must determine which set of answers belongs to the person and which to the computer. This process Turing called an imitation game.

However, this approach provoked distrust and criticism among many scientists and philosophers. The philosopher John Searle strongly criticized the thesis of the defenders of artificial intelligence that consciousness relates to the body just as software relates to the hardware device, which implies that the computer has consciousness and that it has the property of understanding. In particular, Searle criticizes the claims of one of the theorists and developers of artificial intelligence, Roger Schenk, to make a machine that can understand the human language, read stories and answer questions. Schenk created a machine that was able to answer questions seemingly as would a person. Allegedly then, it follows that, first, the machine understands the human language, and secondly, that the computer program is analogous to and explains the ability of the human to understand stories and ask questions.

However, Searle opposes these conclusions with his famous Chinese Room thought experiment:

Suppose he were locked in a room and given a manuscript in Chinese. Searle speaks English, but does not understand Chinese, so for him such a text is a set of meaningless signs. Then he is given a second manuscript in Chinese, as well as the rules for compiling the Chinese characters of the first and second manuscripts in English. Chinese symbols are compared only by formal rules. Now he is given a third Chinese manuscript, as well as the rules for comparing it with the previous Chinese texts in English. The first Chinese text is a story, the second is a manuscript and the third is a question. Symbols that Searle issues in response to the third Chinese text are considered "answers to questions", and instructions written in English are called "program." The author suggests that he seems to have "stuffed his hand" in carrying out instructions to manipulate Chinese characters, and that his answers to the questions of the Chinese text are absolutely indistinguishable from the answers of real Chinese speakers. However, despite the fact that Searle manipulates formal symbols of the Chinese language, he understands nothing of their meaning.

Searle compares his behavior with the behavior of a computer that reliably manipulates symbols, but understand nothing of what they mean. Is it possible to argue that, in this case, the computer program models the human understanding of the language? Also, the activity of the computer cannot explain the principles of human thinking, since neither Searle in the room nor the program in the computer understands, although both effectively perform tasks.

Opponents claim that Searle, as a native speaker of English, even answering questions in English, does the same thing he would do when manipulating Chinese symbols. That is, his thinking, just as the computer's symbol manipulation, is only an operation with conventional signs. However, Searle believes that the computer program has no semantic understanding. "In the case of the Chinese text, I have everything that artificial intelligence can put into me through the program, but I do not understand anything. In the case of the English text, I understand everything, but so far there is no reason to think that my understanding has something in common with computer programs, i.e. computer operations on elements, defined in a purely formal way" [7, p. 380]. Moreover, Searle argues that his ability to speak English probably does not follow strictly formal principles, which is supported by studies in the field of Eco linguistics showing the role of context (pragmatics) and creative component in speech [10, 14]. Also, early in the 20th century the understanding of language and the possibility of intersubjective knowledge was studied in the frameworks of phenomenology, existentialism and hermeneutics [8].

To bolster his position, Searle refers to analysis of the meaning of "understanding". The creators of artificial intelligence insist that there are several meanings or levels of understanding. Searle disagrees and claims that we often attribute understanding metaphorically. For example, the automatic door "understanding" instructions through a photocell is not the same as that of Searle understanding English. The programmed computer of Schenk "understands" the stories no more than the automatic door "understands" the instructions. However, Newell and Simon insist that their computers understand the symbols and are able to know no less than people. Searle simply disagrees, arguing that with them, considering that even if you build a robot with limbs, the ability to see through a camera, and to walk, it would still lack intentionality, because all its motor and perceptual abilities come from the functioning of electrical circuits and programs.

Searle is equally skeptical about creating an artificial intelligence based on neural networks that simulate the brain's neuronal excitation and synaptic activity. Such a machine does not rely on one program, but a set of parallel programs, in the likeness of the human brain. But such machines do not understand the language, they have no intentional states. Manipulation of formal symbols does not confer intentionality. "It's not because I'm able to understand English and have other forms of intentionality," writes Searle, "that I am an instance of a computer program (I'm probably the authority of any number of computer programs), but, as far as we know, that I am an organism of a certain kind with a certain biological structure, and this structure, under certain conditions, is causally capable of producing perception, action, understanding, learning and other intentional phenomena " [7, p. 392].

A computer cannot produce consciousness any more than computer simulation of a fire can burn up a neighbor house, or computer simulation of a shower can make us get wet. However, the reason for this perception of the computer is, in Searle's opinion, confusion of concepts. Thus, the term "information processing" is used both for AI and for the functioning of the human brain. But the information that people process is different from that of the computer, the latter, only manipulates formal symbols. The computer has syntax, but is devoid of semantics: it multiplies "2" by "2" to make "4" but does not understand what these symbols mean. Intentionality is a biological phenomenon, and if we say that mental operations are reducible to computational operations, then we admit that they are independent of biological properties. In this regard, many modern researchers talk about the inapplicability of the technical understanding of information to the study of biological systems and the study of communication processes in nature and society, therefore the code-information paradigm is opposed to the systemic and activity-based approaches [9, 11].

Searle emphasizes: "It does not occur to anyone that we can produce milk and sugar by launching a computer model of formal lactation and photosynthesis sequences, but when it comes to consciousness, many people persistently want to believe in such a miracle because of their deep and strong rooted dualism: the consciousness that they have in mind depends on formal processes and does not depend on very specific material causes – unlike milk and sugar "[7]. A distinctive ability of human consciousness is the ability to comprehend textual meanings, which are expressed implicitly. The computer does not possess such abilities, as it does not possess the human ability to convince. Therefore, to date, according to Searle, the idea of creating human level AI looks utopian.

Another problem related to AI is ontological "fuzziness", expressed in the absence of certainty and the existence of continuous dynamics in many areas, which must be taken into account when modeling the processes of thinking and processing information. The emergence in the early 20th century of research in the fields of stochastic processes and nonlinear dynamics, and development of probability theory, statistical theory, decision theory and led to a revision of an unambiguous, repetitive, mechanistic model of the world based on deterministic laws and contributed to the emergence of a new paradigm of reality based on moderate indeterminism and ontological relativism [1, 3]. Thus, studies of complex nonlinear systems and processes have shown that the same factors acquire different meanings in different reporting systems and change their characteristics depending on time and circumstances. In connection with this, when solving urgent problems related to processing and understanding information or studying the properties (behavior) of a particular system, it is impossible to abstract from various parameters of the relationship, time, and context.

To help deal with uncertainty and unpredictability fuzzy logic and fuzzy set theory emerged. They allow us to work with probabilistic and chaotic processes such as weather or economic change, or when interpretation of multivalued or incomplete information is needed, for example, in the recognition of voice or text.

Classical formal logic is two-valued, "true" or "false". For a long time it was not only considered to underlie rational thinking, but was the basis of digital computing with its binary true/false, yes/no, 0/1 dichotomy. In fuzzy logic, intermediate values between "zero" and "unity" are used in the calculations, which allows us to work with probabilistic, multivalued and variable processes in biological and social systems.

Based on this logic, AI systems were developed that predict the direction of economic processes, or control household appliances. For example, the Japanese financial corporation Yamaichi Securities developed a program that allowed automating the game in the securities market in many ways, and LG released a washing machine based on fuzzy logic. We can also mention studies by Kosco, who developed fuzzy cognitive models and showed the relationship between fuzzy logic and the theory of neural networks, the work of Zemankova and Kandel, on fuzzy database management systems capable of processing fuzzy queries despite incomplete or inaccurate data. There are many other examples of the successful practical implementation of this logical system in a variety of areas [1, 5, and 13].

As a conclusion, summing up the problems in AI, let us turn to some ideas of Martin Heidegger. He notes that modern science differs from ancient. For the Greeks, the main task was to achieve the truth for its own sake. Modern science emphasizes practical applicability. M. Heidegger writes that in science a scholar who serves the topical requirements of highly specialized scientific knowledge comes to replace the scientist-erudite, and this reflects on the way of scientific thinking – the computational way of cognizing the world becomes predominant. Computing thinking is generated by the desire of a person to achieve a certain result in science, technology and everyday life, so it always calculates new promising opportunities. But such thinking, in Heidegger's opinion, is unrestrained and foolish, it "cannot be thought of thinking about the meaning that reigns in everything that exists" [12, p. 104]. A person falls into the power of a formal, digital language and misses the main ontological meanings that cannot be quantified.

In addition to the theoretical issues related to the development of AI and understanding of the nature of thinking, another important question is what will be in the future with human intelligence tied to the computer? Our inner world is changing rapidly. A person has never had such easy access to the world's information database. It's hard to even imagine what will happen to the human intellect in the future. It is already clear that everyone who has mastered the skills of working with computer equipment becomes dependent on it. Either a dependence on computer games and social networks, especially in children and young people, or viewing diverse, vivid multimedia information. For a modern scientist, a computer is also a best friend, yielding information and knowledge. So who creates whom? Does human intelligence become artificial?

Bibliography

1. V.Vedernikov. Fuzzy-multiple modeling in the analysis and forecasting of economic phenomena and processes: the historical aspect // Problems of the modern economy, 2006. No. (17). URL: www.hse.ru/data/868/318/1234/Vedernikov.doc

2. Gerovich VA Dynamics of research programs in the field of artificial intelligence: the thesis ... Candidate of Philosophical Sciences: 09.00.08 / AN USSR. Institute of Philosophy. – Moscow, 1991.

3. Glebov RE, Sukhoverkhov AV Relativism, multi-valued logic and information theory // Information technologies in self-realization of personality. – Krasnodar: KubGAU, 2012, – S. 198-206.

4. Dreyfus H., Dreyfus S. Creation of consciousness vs. Modeling of the brain: artificial intelligence returned to the branch point // Analytical philosophy. Formation and development (anthology). Per.s English, German. M.: – "House of the intellectual book", "Progress-tradition", 1998.

5. Kalinnikov V.A. The use of multi-valued logic in digital technology. // Devices and technics of experiment. 2006. № 6. With. 5-14.

6. Luger JF Artificial intelligence: strategies and methods for solving complex problems, 4th edition .: Trans. With the English. – M.: Publishing house "Williams", 2003.

7. Searle D. Consciousness, the brain and programs. // Analytical philosophy. Formation and development (anthology). Per.s English, German. M .: – "House of Intellectual Book", "Progress Tradition", 1998.

8. Spasova N. E. The emergence of the problem of intersubjectivity in the New European philosophy and its transformation in modern philosophy. Krasnodar, 2010.

9. Sukhoverkhov AV Methodology of "situational" and "contextual" knowledge and training // Innovative technologies in the educational process as a resource to improve the level of training of specialists. The collection of theses of the interfaculty educational-methodical conference. – Krasnodar: KubGAU, 2013. – P. 231-233.

10. Sukhoverkhov AV Current trends in the development of ecolinguistics // Language and Culture, 2014. No. 3 (27). – C. 166-175.

11. Sukhoverkhov AV Language, communication, inheritance: nonrepresentative models // Science and education as the basis for self-realization of personality. Krasnodar: KubGAU, 2014. Issue 3. – P. 88-92.

12. Heidegger M. Detachment // Conversation on the country road. Collection. Trans. With him. Ed. A.L. Dobrokhotova. – M .: Higher education. Shk., 1991.

13. McNeill D., Freiberger P. Fuzzy Logic: The Discovery of a Revolutionary Computer Technology and How It is Changing Our World. N.Y., 1993.

14. van Lier L. From input to affordability: Social-interactive learning from an ecological perspective. In: J. P. Lantolf (ed.) Sociocultural Theory and Second Language Learning. Oxford: Oxford University Press. 2000. P.155-177.

A.V. Sukhoverkhov
Candidate of Philosophy and Associate Professor
Department of Philosophy
Kuban State Agrarian University

A.V. Spasova
Candidate of Philosophy and Associate Professor
Department of Philosophy
Kuban State Agrarian University

Email: ksau2009@gmail.com, natalia.spasova@mail.ru