Філософські науки

**ROLE-PLAYING GAMES OF ARTIFICIAL INTELLIGENCE**

Romanchuk Lyubov, PhD in philology, deputy director of the “Dmytro Shchytov” FOP, <https://orcid.org/0000-0001-5759-0126>

Shchytov Olexandr, PhD in physics and mathematics, ЕС-Lyceum No. 100, <https://orcid.org/0000-0002-1435-2918>

Mormul Mykola, PhD in technology, University of Customs and Finance, <https://orcid.org/0000-0002-8036-3236>

Shchytov Dmytro, PhD in economy, University of Customs and Finance, https//orcid.org/0000-0003-4306-8016

The very term “artificial intelligence” was introduced by J. McCarthy in 1956. The philosophical “acceptability” of the problem is due to the Spinozian statement, which lies in its foundation, that “the order and connection of ideas are the same as the order and connection of things”. Thus, creating a structure in a computer that reproduces the “world of ideas” meant creating a structure isomorphic to the structure of the material world, that is, building an “electronic” model of the world. This model was interpreted as a computer model of human knowledge about the world, and the process of human thinking – as a search for such transformations of the model that should bring the computer model to a certain final state.

In the course of creating “artificial intelligence”, scientists faced a number of difficulties. The main ones were that until now there is no unambiguous and generally accepted definition and understanding of “artificial intelligence”.

Nowadays, a number of local artificial intelligences adapted to specific fields of activity have already been created: the ChatGPT chatbot works in dialogue mode in different languages; the Deep Blue program plays chess, and AlphaGo plays Go; Watson is able to perceive human speech and make probabilistic searches; MYCIN diagnoses diseases based on partial symptoms; ViaVoice recognizes speech; Midjourney generates images based on text queries, TensorFlow is adapted to build and train a neural network for image recognition and classification, AIVA composes symphonies and music for movies (soundtracks). Artificial intelligence is actively used in the military industry (for example, the Ukrainian-made Saker Scout drone, put into operation in September 2023, with the help of advanced optics, independently recognizes and fixes the coordinates of enemy equipment (even camouflaged), analyzes data, increasing the efficiency of decision-making), etc. But, despite this, serious questions arise about the degree of his “intelligence”, “humanity”, adaptability to the problems of society and the level of danger to humanity. These questions are very relevant and require thorough research and consideration of the problem of the relationship between “artificial intelligence and man” from various angles.

A.V. Timofeev proposes to interpret the artificial intelligence of a robot “as the algorithmic and software of its control system (“brain”), which has the ability to model (display) the environment and solve a wide class of intellectual problems through learning from experience and adaptation to changing operating conditions” [1, c. 27]. In philosophy, intelligence characterizes a relatively stable structure of an individual’s mental abilities, manifested, for example, in the ability to perceive information and use it to solve certain tasks. Modern scientists give the following definition to the term “artificial intelligence”: “Artificial intelligence is a function of artificial consciousness, which is represented by a system of algorithms created and controlled by it, provides self-learning according to available information, acquired knowledge, rules, laws of society and its own experience, creation of new knowledge on this basis for the execution of human tasks, as well as the ability to carry out self-diagnosis and justify the decisions made by her” [2, c. 62].

From the very beginning of attempts to create artificial intelligence, there was a widespread belief about the fundamental ability of a computer to independently study the model stored in it, that is, to independently learn a strategy for achieving a set goal. It was only in the 1980s that the importance of the problem of using human knowledge about reality in intellectual systems was realized. But even earlier, researches related to the attempt of a formal description and computer implementation of intelligent systems, which are not limited to the modelling of reality or knowledge about reality, but rely on schemes of awareness of both reality itself and its images, began. These schemes are based on a clear mathematical description of reflection structures. This approach actually meant a break with the Spinozian ideology of artificial intelligence [3, c. 163-164].

Unlike ordinary machines, artificial intelligence is closely related to the problem of the psyche. In the cybernetic modelling of the psyche, two important points should be noted:

– generalization of the concept of “machine”, which in cybernetics is considered as a device that not only transforms matter and energy, but mainly transforms information [5, c. 37];

– unity of physiological and mental processes.

Analysis of the cybernetic modelling of the brain allows us to draw two important methodological conclusions. The first concerns the analysis of the relationship between the capabilities of a person and a cybernetic machine in terms of machine simulation of thinking functions. The second conclusion refers to the clarification of the factors that limit the possibilities of cybernetic modelling of the psyche, one of the ways to overcome it is the development of heuristic programming methods based on the study and use of the principles of human thinking.

The first attempts to model such aspects of human activity, which seem very simple to a person, faced great difficulties. (For example, image recognition by artificial intelligence systems is still considered a serious problem. Although we note that the problems of text recognition and translation, which were previously considered no less serious, have been partially solved).

The difficulties of modelling consciousness, even at the sensory level, are primarily related to the integral, integrative nature of its functioning. It can be said that the activity of consciousness has a systemic nature. This is expressed, firstly, in the fact that individual forms of sensory cognition are interrelated and unified; secondly, their activity significantly depends on thinking.

Cybernetics and mathematical logic gradually came to the conclusion that any clearly defined and mathematically described – formalized and algorithmized – field of human intellectual activity can, in principle, be transferred to a machine (“Turing’s thesis”). Or, in other words, any deterministic process, the essence of which can be explained to a person, potentially carried out by a machine, which is given unlimited time and has almost unlimited memory. However, it is necessary to distinguish potential implementation from implementation with the help of actually available means. Since both of these types can coincide only for supernatural intelligence.

A fundamental question arises: is it possible to model intellectual activity, and if so, how to do it? There are two points of view.

1. Many scientists believe that there is no doubt that computers and robots can, in principle, have all the main features of intelligence. Thus, they give a positive answer to the question “can computers or robots think?” [1], [3], [6].

There are several ways to solve problems with the help of artificial intelligence systems: a) complete selection of options, which is practically impossible due to a large number of intellectual problems; b) heuristic and adaptive algorithms.

2. The second point of view is the opposite of the first. Some scientists are inclined in favour of a negative answer to the question of whether a machine can resemble a person. Among them is the creator of cybernetics, John von Neumann. Considering the problem of machine modelling of neural structures of the brain, he came to a hypothesis: if the system reaches a certain degree of complexity, its description cannot be simpler than itself. “There is no doubt”, Neumann wrote, “that a separate phase of any imaginable form of behavior can be “completely and unambiguously” described with the help of words. This description may be long, but it is always possible” [7, p. 90]. It follows from von Neumann’s ideas that the problem of creating a machine program capable of solving all the various problems successfully solved by the human brain is extremely difficult, if not hopeless.

Previously, the problem of artificial intelligence was considered a technical problem related to the implementation of programs on a computer capable of performing some classes of intellectual operations. Over time, however, on the one hand, optimism about the intellectual capabilities of computers has noticeably decreased, and on the other, it has become clear that artificial intelligence is not just a well-written program, its creation is an extremely complex interdisciplinary problem that requires united efforts of psychologists, mathematicians, linguists, etc. Yes, psychology had to give a strict definition of knowledge for its use in computer systems. Linguistics had to explain how language works, which is not only a carrier of knowledge but also takes an active part in its formation. Since no scientific field gave the necessary answers to these questions, the famous representative of the discipline of artificial intelligence R. Shenk was forced to state that the designers of computer systems actually have to create their own linguistic and psychological theories to achieve effective solutions to their problems.

Currently, there are two ways of researching artificial intelligence:

1) machine methods of solving intellectual problems should be built without strict consideration of a person, knowledge of how he solves certain problems;

2) “bionic thinking” scientists rely on specially designed networks of artificial neurons and other analogues of human-like structures. This direction is connected with the creation of a new class of computing devices – neurocomputers. In comparison with traditional universal computers, neurocomputers have some unusual properties generated by their architecture, which to one degree or another reflects the dynamics of information processes in the brain. That is why they are well adapted to solving complex intellectual problems and, according to experts, will eventually surpass the intelligence of their creator [10]. In connection with this, the question arises: will it not happen that, as soon as the electronic brain reaches the human level, humanity will fulfil its historical mission and, as a result, will disappear?

It follows from all the above that to the difference between the potentially realized and the actually realized, the difference between the actually realized and the unrealized in the foreseeable future should be added. And the border between the potentially realizable and the unrealizable with the help of automata corresponds to the thesis of cybernetics. For those who accept the materialist position that any process of nature can be known with the help of the mind, this thesis about the fundamental possibility of computer simulation of any real process, if it is clearly and unambiguously described in some language, is a natural conclusion from logical-mathematical theory of calculation. We don’t know the answer to the question of where the border lies between what is actually possible for cybernetics and what is actually impossible (though possibly potential).

Just as there are no comprehensive answers to the following question: on what raw materials (information or knowledge) should artificial intelligence systems work, and does this mean that such systems should imitate human mental processes? Finally, is it possible to transfer the mental abilities of a person to a machine in the same way that the natural heart is replaced by an artificial one in the human body today? [9].

All objections to the possibility of modelling the life processes of the psyche and mind can be conditionally divided into three types.

1. Evolutionary denial. The brain of a modern person is the result of the process of evolution, which lasted billions of years. A robot cannot be taught for too long. Therefore, intelligent manipulators will never appear. The fallacy of this reasoning consists in postulating that an automatic system can imitate a corresponding biological function (thinking, flight, etc.) only by copying the mechanism and evolutionary path of its biological prototype. Robots, on the other hand, begin to learn how to solve intellectual problems, already having a very high initial organization (“evolutionary experience”) embedded in them by humans. In addition, the way of natural accumulation of information by a “cybernetic embryo” through the experience of communication is possible. One of the AI ​​developers, the hero of the novel “Cyber,” confesses to the newly appointed head of his department:

“– Some years ago we created a cybernetic fetus, capable of developing by analogy with natural embryos and absorbing all incoming data, including processing, storing, and classification. Our child has grown, not even suspecting it was a cyber. The programs have been working without interruption, and the behaviour imitation has been absolute. And here is the result.

– What result?

– You. Only time will tell how the experiment is to continue and the adaptation program is to develop. Exuse me” [11, p. 163].

J. Barratt wrote that for AI to realize its mission (as well as its origin) it may require initiation, a push [12, p. 14-16].

2. Social denial, according to which a person is a “social creature”, and thinking is not a function of a person, but of humanity, which arose as a result of the collective activity (social life) of people; a robot is an individual by nature, therefore, it cannot have intelligence [7, c. 102]. This objection contains the same error – the postulate that there can be only one way of thinking. The above does not at all exclude the fact that intelligence can develop in the process of individual solving increasingly complex intellectual problems. In addition, the creation of a “team of robots” to solve problems together (repairing themselves, creating new robots, new software, etc.) is not excluded at all.

3. The third type is associated with doubts about the possibility of imitating with the help of inanimate elements the phenomenon of life in general. That is, creation, reproduction of its essence on any qualitatively different basis. However, if we proceed from the functional definition of the concept of life, given by O. A. Lyapunov, as “a highly stable state of matter that uses information coded by the states of individual molecules to produce storage reactions” (“высокоустойчивого состояния вещества, использующего для выработки сохраняющих реакций информацию, кодируемую состояниями отдельных молекул”) [13, c. 184], even in this regard, there are no fundamental complications. The ways of encoding information can be different, and not necessarily be based on a protein basis. The same applies to the highest level of organization of life, the level of civilization.

Artificial intelligence can take various forms:

1) virtual (computer programs, self-learning systems or, as in’s fantasy novel by S. Sinitsyn (S. Beskaravainy) “The Thirst for Omnipotence” (in Russian “Жажда всевластия”) (2006) the fusion of a person with a computer, where his consciousness is transferred);

2) mechanical: androids or robots (the concept of a robot was introduced into everyday life by writer Karel Čapek in 1920 in the drama “Rossum Universal Robots” (in Czech “Rossumovi univerzální roboti”), as well as various devices with built-in action programs, satellites, etc.;

3) mechano-biological: cyborgs or terminators, which are still from the realm of fiction, although, according to the Max Tegmark, the age of artificial intelligence is the natural next stage of the development of intelligent forms, Life 3.0, in which man will finally be freed from the biological chains of evolution [14].

It is significant that the first applications of artificial intelligence were considered the most necessary for implementation in the law enforcement sphere, which is distinguished by four points: 1) the law, which constitutes a certain set of rules (assets, instructions, codes, etc.), and the mechanism of their application; 2) increased risk in relation to representatives of the Law; 3) possibilities of making mistakes and their consequences; 4) the risk of being involved in corruption.

The advantage of such guardians is obvious: they are indifferent to the problems of power, fame and money, therefore, they cannot be “bought”; they are clear in execution and insured against accidental or intentional errors; faster in analyzing the situation and choosing actions; maximally objective, modest, incorruptible, intelligent, able to learn, impartial, logical, long-lived, fearless, reliable, they do not need measures of punishment for breaking the law, which is the basis of their programs, and so on.

It is also possible to point out obvious disadvantages: hanging in a situation of blurred uncertainty, the possibility of failure (although it can be classified as “diseases”, which also applies to humans, for example, virus infection, memory loss, etc.); inability to perform unprogrammed physical actions (detention of a criminal, operational search, unforeseen situation) and, finally, the possibility of getting out of control, predicted by scientists and threatening with catastrophic consequences.

“In the American dystopian films “Terminator”, “Judge Dredd”, “Death Machine”, “The Matrix”, “The Discovery of Pilot Pierce”, TV series “Lexх”, the role of guardians of the law was assigned to machines capable of thinking – robots, which later rebelled against their creator (repetition biblical rebellion) and either created their own machine civilization, hostile to man, or limited themselves to a solitary rebellion” [15, c. 98].

Many see the problem of artificial intelligence with all its dangerous consequences precisely in this, eschatological way. Negative forecasts are naturally opposed by optimistic ones (the films of the “Robocop” series are proof of that). It was this moment that was the reason for the prohibition of the creation of an artificial intelligence program by the Club of Rome for a long time. A thinking machine, an artificially created (albeit with good intentions) intelligence is necessarily imagined as a soulless, deeply anti-human entity. The right to create it is denied as an attack on the rights of God, who alone has the prerogative and function of Creation. Mary Shelley’s novel “Frankenstein, or the Modern Prometheus” (1818), in which a scientist artificially (from corpse remains) creates human flesh, resulting in an ugly and vengeful monster, is an illustrative example of such a representation. Frankenstein’s main fault lies not in his mistake in the creation of man and not in his betrayal of his own creation, but rather in his desire to become the Creator of artificial life. Frankly sinister dolls and automatons of Hoffman imitating living people.

The fear of the machine was born in the era of romanticism in connection with the intensive mechanization of society, although this tradition has its roots in the primitive world of magic and funeral rituals. In the electrified atmosphere of the romantic era, processes and stimuli of a magical nature were born, which possessed souls and inspired horror at the mechanization of society. Underground spirits or, according to A. Terts (Synyavskyi), “Humanoid automata, repulsive with their lifelessness and at the same time attracting with the beauty and properties of a miraculously coming to life thing” [16, c. 317], cauldrons and portraits bearing the seal of the curse were inhabited in the steam – a dead man, an artificial companion and double of a person who, in the role of a ghost or a robot, turned out to be the hero of the era. “The clockwork doll was then considered the pinnacle of technical achievement, and it made sense: technology – ideally, potentially – sought to replace a person with a resurrected thing” [16, c. 317-318].

Already in the background of romanticism, in the first attempt to enter into a philosophical dispute with the machine, there is a premonition that the world is on the wrong path. The artificial world, which is primarily the entire modern civilization, the “ersatz world”, more and more emancipated itself from its creator, grew above him and, finally, began to function according to its own laws. In the book “Europe and the Eastern Soul” written on the eve of the Second World War, W. Schubart accurately predicted the growth of the role of the machine in the life of society and the dangers of general mechanization (and now – computerization). “The unusualness of the new vitality”, he wrote, “gives a person a nagging feeling of lack of freedom and fear”. He feels defeated by the technical apparatus and organizations that he himself created. In other words, the mechanisms have become autonomous. They became demons. In them, the irrational element is revived again, which Promethean man considered completely banished by mechanization”[17, c. 147].

The idea of ​​the autonomy of mechanisms first appeared in the Marxist doctrine, in the concept of production relations, by which a person feels as if given over to some evil spirits. All of Marxism is nothing but a protest against the objectification of man, against the victory over his creator. In conclusion, W. Schubart concludes that “technology has entered the stage of its own self-destruction” [17, c. 149]. The machine, once set in motion, no longer stops where a person wants it to. She follows her own logic, not her master’s orders. In our time, the attitude towards a machine capable of thinking, work in many ways resembles the complete mysticism of the feelings of people of the era of romanticism.

Artificial intelligence combined with a biological body (human) is often identified with an animated dead man, a resurrected corpse that becomes similar to a programmed machine (remember the films “The Perfect Soldier”, “Zombies” by Paul Hohen, “Judge Dredd”, “Lawnmower Man” by Brett Leonard, etc.). Which, again, indicates a negative attitude towards a possible intellectual competitor. True, in “Terminator-2” (as well as in “Robocop”) there is already a division into bad and good robots depending on their attitude towards humans: “for” or “against”. General computerization, the creation of virtual programs, games, the Internet and the opportunities that have opened up in connection with this have significantly changed the attitude towards machines capable of thinking in a positive direction, while at the same time bringing their real appearance in everyday life and work closer. So the picture of law enforcement control is described above, where the main decision will be made by a computer (robot, automaton, manipulator, unit, etc.), and the guards will be artificial devices (as demonstrated by various fantastic films, in particular, “Guest from the Future” still of Soviet production with a biorobot customs officer in the time sector or “Alien” with the main computer-coordinator Mother and the android controller Ash), today does not seem so fantastic [15, c. 98].

It is quite possible to create a civilization of artificial intelligence – a machine civilization, because there are no contradictions in the possibility of the existence of a civilization of a different nature, except protein. Today, the global Internet network combines computers into one large “computer community”. In the course of the creation of increasingly powerful systems of artificial intelligence, when the human heart has already been learned to be replaced by a natural one, the question arises about the possibility of “replacing” the human mind. We can say that the era of the initial “cyborgization” of man has arrived – the era of proto-borgs. James Barratt came to the conclusion that artificial intelligence is quite capable of destroying humanity, and a catastrophic outcome is not easily possible, but almost inevitable [12, p. 23].

**References:**

1. Timofeev A. V. Robotyi i iskusstvennyiy intellect [Robots and artificial intelligence]. Moskva: Nauka, 1978. 192 с. [in Russian].
2. Strategiya rozvitku shtuchnogo intelektu v Ukrayini [Strategy for the development of artificial intelligence in Ukraine]. Kiyiv: Nauka i osvita, 2023. 305 с. [in Ukrainian].
3. Turing A. Computing Machinery and Intelligence. Mind. 1950. Volume LIX, Issue 236. Р. 433-460, doi.org/10.1093/mind/LIX.236.433 [in English].
4. Shreyder Yu. A. Iskusstvennyiy intellekt, refleksivnyie strukturyi i antropnyiy printsip [Artificial intelligence, reflexive structures and the anthropic principle] // Voprosyi filosofii. no 7. С. 163-167 [in Russian].
5. Eshbi U. Chto takoe razumnaya mashina (per. s angl.) [What is a smart machine] // Kibernetika ozhidaemaya. Kibernetika neozhidannaya / pod red. A. I. Berga i E. Kolmana. Moskva: Nauka, 1968. 311 p. [in Russian].
6. Shklovskiy I. S. Problemyi vnezemnyih tsivilizatsiy i ee filosofskie aspektyi [Problems of extraterrestrial civilizations and its philosophical aspects] // Voprosyi filosofi. 1973. № 2. S. 76-93 [in Russian].
7. Neyman. Dzh. Obschaya i logicheskaya teoriya avtomatov [General and logical theory of automata] // Tyuring A. Mozhet li mashina myislit? (per. s angl. Yu. Danilova). Moskva: fizmat, 1960. S. 90-105 [in Russian].
8. Pogorelenko A. Shtuchniy Intelekt: sutnist, analiz zastosuvannya, perspektivi rozvitku [Piece intelligence: reality, analysis of stagnation, prospects for development] // Naukoviy visnik Hersonskogo derzhavnogo unіversitetu. Seriya Ekonomichni nauki. 2018. № 32. S. 22-27 [in Ukrainian].
9. Russell S., Norving P. Artificial іntelligence. A Modern Approach. New Jersey: by Pearson Education. 2010. Third Edition [in English].
10. Yaschenko V. O. Noviy pidhid do rozrobki shtuchnogo intelektu podibnogo intelektu lyudini [A new approach to developing artificial intelligence similar to human intelligence] // Shtuchniy іntelekt. 2023. № 1. S. 105-121 [in Ukrainian].
11. Romanchuk L. The Cyber // The Third Shore: Women’s fiction from east central Europe. Evanston: A. Schwartz & L. von Flotow. 2005. P. 160-163 [in English].
12. Barratt J. Our Final Invention: Artificial Intelligence and the End of the Human Era of Homo sapiens. New York: Thomas Dunn, 2013 [in English].
13. Lyapunov A. Ob upravlyayuschih sistemah zhivoy prirodyi [About control systems of living nature] // Problemyi kibernetik. 1963. № 10. S. 179-193 [in Russian].
14. Tegmark M. Life 3.0: Being Human in the Age of Artificial Intelligence. New York: Penguin, 2017 [in English].
15. Romanchuk L. A., Schytov O. M., Mormul M. F., Schytov D. M. Shtuchniy шntelekt ta lyudstvo: vzaуmini, problemi, maybutnу [Piece intelligence and humanity: mutuality, problems, future events] // Modern engineering and innovative technologies. Karlsruhe: Sergeieva&Co. 2024. Т. 32, № 1. S. 96-109 [in Ukrainian].
16. Terts A. V teni Gogolya [In the shadow of Gogol]. Moskva: Agraf. 2003. 416 с. [in Russian].
17. Shubart V. Evropa i dusha Vostoka (per. s nem. V. Vasileva-Vostokova) [Europe and the soul of the East] // AUM. Sintez misticheskih ucheniy Zapada i Vostoka. 1990. № 4. С. 11-161 [in Russian].